RESEARCH DEGRESS WITH PLYMOUTH UNIVERSITY

OPTIMAL RESPONSE IN DECISION MAKING: AN EXPERIMENTAL INVESTIGATION OF DECISION STRATEGIES

by

VALERIO BISCIONE

A thesis submitted to Plymouth University
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

Centre for Robotics and Neural Systems.

Plymouth University

September 2016

Acknowledgements

This thesis would not have been possible without the support, the wisdom, and the guidance of many scientists that I had the pleasure of meeting during my staying at Plymouth University.

First and foremost, I wish to thank my supervisor Professor Chris Harris, who gave me invaluable support since the first day I began my journey as a researcher. His guidance, encouragement, and enthusiasm were the scientific fuel that propelled my research during the past four years. I am also grateful to my second supervisor, Professor Roman Borysiuk. His seminar on advanced mathematics, which he taught with passion and joy, helped me get a better grasp on more sophisticated topics in cognitive modelling. I would also like to thank my thesis committee for taking the time to read this work and for providing feedback on the research I have conducted.

On a more personal note, I am sincerely grateful to the large number of people that made my time in Plymouth unforgettable. I am thankful to Debora Zanatto, Alexandre Antunes, Frederico Klein, François Blake, Emmanuel Senft, Mina Marmpena, Martina Fiori, Karla Štěpánová**,** Ricardo de Ajambuja for making the office such an enjoyable environment to work in; Giuseppe Filippone and Davide Spataro, whose precious advice on numerical modelling were balanced by long breaks for table tennis; Valentina Valassina, for her fashionable advice; Inge Tamburrino, for organizing adventurous trips; Cristina Gnisci, who cheered us up with her delicious food; and to the three of them, as a team: our time together, in Plymouth and outside Plymouth, will always have a special place in my memory. Exceptional thanks to the Italian Alpine Ibex team: Vincenzo Motta, Luca Stirnimann, Riccardo Polvara, Andrea Ferrario, and the Saltash friends. Through countless coffees, meetings, lunches and dinners, they made me feel surrounded by family even hundreds of kilometres away from home. I wish our bond will hold for the next years, even when we will be scattered all around the world. Part of the Alpine Ibex team, but deserving some special thanks on their own, are my two dear friends and roommates: Giovanni Sirio Carmantini and Massimiliano Patacchiola, two of the most brilliant minds I had the

pleasure of meeting during this endeavour in Plymouth. Your devotion to science and research will stay in my heart as an example that I will never forget.

Some relationships are meant to last through space and time: such is my friendship with Lorenzo De Donato and Daniele Santorelli, whom I deeply thank for always being available when my mind needed to wander away from work.

My mother Tiziana, my father Angelo and my brother Marcello have been extremely helpful all throughout this journey: they gave me affection and unconditional encouragement, especially during the darkest times. They have provided me with the best emotional support I could have ever asked for.

Finally, I owe a very special thank you to my other half Bahar İrfan. Your patience, gracefulness, and kindness, brightened up my days in Plymouth and helped me maintain a positive perspective in life. May our days together be long, warm, and full of joy.

Author's declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

Work submitted for this research degree at the Plymouth University has not formed part of any other degree either at Plymouth University or at another establishment.

Signed:

Date:

Publications:

Biscione, V., Harris, C. M. (2015) Investigating decision rules with a new experimental design: the EXACT paradigm, *Frontiers in Behavioral Neuroscience. 9:288. doi: 10.3389/fnbeh.2015.00288*.

Harris, C. M., Waddington, J., Biscione, V., & Manzi, S. (2014) Manual choice reaction times in the rate-domain, *Frontiers in Human Neuroscience. 8:418. doi: 10.3389/fnhum.2014.00418*

Presentation and Conference Attended:

Biscione, V., Harris, C. M. (2015) A new approach in studying Decision-Making: the EXogenous ACcumulation Task (EXACT). *25th Conference in Subjective Probability, Utility and Decision Making.* Budapest, Hungary. *(Oral Presentation).*

Biscione, V., Harris, C. M. (2015) A New Paradigm for Investigating Human Decision Strategy, *Proceeding of the International Psychological Applications Conference and Trends.* May. Ljubljana, Slovenia. *(Poster Presentation)*

Biscione,V., Harris, C. M. (2015) Connection Piéron's Law, the Foreperiod Effect and Distribution Shapes in a Simple Reaction Time Task*, Proceeding of the International Psychological Applications Conference and Trends*. May. Ljubliana, Slovenia*. (Oral Presentation)*

Biscione, V., Harris, C. M. (2015) How to investigate decision-making in humans? *Plymouth Postgraduate Conference Series.* Plymouth. *(Oral Presentation)*

Biscione, V., Harris, C. M. (2014) Rate Domain Distributions in Simple Reaction Time Task. *First Decision Making Conference*. Bristol. *(Oral Presentation)*

Biscione, V., Harris, C. M. (2014) Decision Making and Rate of Reward. *The 6th Annual Plymouth University School of Psychology Conference*. Plymouth. *(Oral Presentation)*

Word count of main body of thesis: 64,563

Abstract

OPTIMAL RESPONSE IN DECISION MAKING: AN EXPERIMENTAL INVESTIGATION OF DECISION STRATEGIES

Valerio Biscione.

A decision process can be conceptually separated into a perceptual process and a decision strategy. The former includes all the different mechanisms that contribute to accumulate information relevant to the decision, whereas the decision strategy determines when enough information has been accumulated and a decision can be taken. Although perceptual processes have been extensively investigated in the last decades, decision strategies have received comparatively little attention. The main aim of this work is to fill this gap by analysing four decision strategies with two different experimental paradigms. We also focus on ancillary decision-making topics, such as the effect of stimulus intensity, foreperiod duration, payoff manipulation, and the response distributions in the rate domain.

We initially performed a qualitative analysis of decision strategies by using a classic reaction time tasks on human participants while assuming the Drift Diffusion Model, one of the many models used for simple and fast decisions, as the perceptual process. We found that increasing the time of the trial does not have a relevant effect on the response, which is in contrast with some of the decision rules considered here. However, this approach is limited by the implicit assumption of a perceptual model that would result in different prediction for the decision strategies.

We suggest the use of a different experimental design, called the EXACT Paradigm, which allows us to analyse decision strategies without having to assume any perceptual process. We tested the feasibility of such approach and applied it to several experimental studies, including a direct comparison with a classic reaction time task. Overall, two of the four decision strategies (modified Reward Rate and

Reward/Accuracy) appeared to model the data satisfactorily. We discuss several ways in which the EXACT Paradigm can be used for expanding our knowledge in the field of decision-making.

Contents

List of Figures

List of Tables

Table of Abbreviations

Miscellaneous	
AIC	Akaike Information Criterion
BIC	Bayes Information Criterion
DDM	Drift Diffusion Model
df	Degrees of freedom
EXACT	EXogenous ACcumulation Task
FP	Foreperiod
RWI	Response-to-Warning Signal Interval
RSI	Response-to-Stimulus Interval
RT	Reaction Times
LATER	Linear Approach to Threshold with Ergodic Rate
mFP	Mean Foreperiod
mRWI	Mean Response-to-Warning-Signal Interval
NAFC	N Alternative Forced Choice (e.g. 2AFC)
STARS	Sequential t-test Analysis of Regime Shifts

Pure DDM	
$\boldsymbol{d}$	Drift rate
$\boldsymbol{a}$	Threshold separation (distance between two thresholds for Choice RT, or between threshold and starting point for Simple RT)
$\boldsymbol{\tau}$	Sensory-motor component
$\boldsymbol{s}$	Noise coefficient (fixed at 1)
$\boldsymbol{b}$	Bias Parameter (set equal to 0.5: no bias)
$\boldsymbol{z}$	Starting point ($z = a \cdot b$)

Extended DDM (when different from Pure DDM)	
$\boldsymbol{\tau}$	Mean Sensory-motor component (non-decision component of RT, $RT = DT + \tau$)
$\boldsymbol{s_t}$	Range of the Uniform Distribution for the variability in τ
$\boldsymbol{s_z}$	Range of the Uniform Distribution for the variability in the starting point
$\boldsymbol{d}$	Mean drift rate (Normally distributed across trials)
$\boldsymbol{\eta}$	Standard Deviation of the drift rate across trials

Decision Rules	
$\boldsymbol{ACC(DT)}$	Speed-accuracy function
$\boldsymbol{RR}$	Reward Rate
$\boldsymbol{RR_m}$	Reward Rate Modified
$\boldsymbol{RA}$	Reward/Accuracy
$\boldsymbol{BR}$	Bayes Risk
$\boldsymbol{\lambda}$	Speed of increase accuracy in time
$\boldsymbol{\alpha}$	Starting point
$\boldsymbol{q}$	Emphasis on Accuracy (punishment/reward ratio for RR_m)
$\boldsymbol{D}$	Delay across trials (RSI, RWI, FP)
$\boldsymbol{D_{tot}}$	Total trial time that is not decision time ($D_{tot} = D + \tau$)
$\boldsymbol{DT}$	Decision time ($DT = RT - \tau$)

Chapter 1

Introduction

As goal-oriented living organisms, competing with other living organisms in an environment with limited resources, we must face decisions in order to obtain the reward we long. We not only must be able to extrapolate information coming from a myriad of sensory sources, but we need to decide when enough information has been accumulated, and thus commit to a certain decision. As this competition has been going on since the beginning of life, it seems reasonable that we have gotten reasonably good at it. How good, exactly, is a question that scientists have tried to answer for the last hundred and half years, by measuring the speed and the accuracy of human responses for simple tasks. This investigation can be tracked back to 1849, when Helmholtz measured the speed at which a signal is carried along a nerve fibre (around 24.6-38.4 meters per second); around 1865, partly inspired by Helmholtz work, a Dutch physiologist named F. C. Donders became interested in measuring the speed of mental processes by using electric shocks to the right or left foot of human subjects. Participants were asked to press a key corresponding to the foot that had received the shock. In one condition participants knew in advance which foot was going to receive the electric shock, in the other they did not. In today's terminology, this would correspond to a simple and a choice reaction time task. Donders found a difference between the two conditions of around 60 milliseconds, establishing for the first time a central tenant of cognitive psychology: simple reaction times are faster than choice reaction times. Based on Helmholtz and Donders contributions, W. M. Wundt established in 1875 the first formal laboratory for psychological research at the University of Leipzig, marking psychology an independent field of study. Much of the research in Wundt's laboratory was concerned in confirming Donders' findings on reaction time and in establishing that mental processes are measurable and quantifiable: the field of experimental psychology and mental chronometry was born.

While these and other researchers were collecting data about the speed of several mental processes, others were interested in how accurate these mental processes are, and the relationship between their speed and accuracy. This relationship was investigated by Woodworth (1899) and, around the same time, by Martin and Müeller (1899), for simple obligatory movements. This brought about the first experimental finding that the accuracy of a response increases with decreasing speed. Subsequently, the study of Henmon (1911), in a discrimination task in which participants were asked to select the longest line amongst two possible options, found an orderly relation between the time and the accuracy of response. A decade later this result took the name of "speed-accuracy relation" (Garrett, 1922; for a review, see Heitz, 2014), before being then forgotten for the next three decades. At this point, the field of decision making started working on formulating a mathematical relationship between RT and accuracy, in the form of sequential sampling of information (Edwards, 1954; Audley, 1957; 1958; Stone, 1960), using the sequential probability ratio test developed by Bernard (1946) and Wald (1948). In particular, Stone combined the speed-accuracy trade-off with the mathematics and optimality of the sequential probability ratio test developing a model known as *random walk* which made specific and testable predictions about response distributions. This model became extremely popular, mainly thanks to its ability to explain within the same framework both the speed and the accuracy of a response, whereas previous models could take into account only one or the other feature. From that point on, several variants of sequential sampling emerged (as discussed in **Section 2.2**). This corresponded to a massive improvement in our understanding of how the decision making process works in humans. However, these models, by themselves, did not offer a complete explanation of the decision making process: they provided us with a relationship between speed and accuracy, but they did not tell us how good we are in making a decision, because they are agnostic regarding what constitute a good decision. In other words: we know that some speed of responses will most likely correspond to some value of accuracy, but why do participants choose that speed or that accuracy? What rules are they using? And how good they are in following them? How these rules depend on the experimental conditions adopted?

This problem has generally been investigated as an optimality problem. The concept of optimality in the study of human mind was not new. Philosophers and psychologist have long discussed the

possibility that people's reasoning followed the rules of formal logic or probability theory (Braine, 1978; Holland, et al., 1986; Edwards, Lindman, & Savage, 1963; Kahneman & Tversky, 1973).

However, we believe that this approach may lead researchers astray. Before investigating how good people optimize in reaction time tasks, we need to understand what does optimization mean in this context. Only by understanding what the *objective function* of the optimization process is we can then investigate how close human participants can get to it. Using a less formal terminology, we will refer to the objective function (the quantity the participants are trying to maximize) as the *decision rule* of the decision process.

This problem was deeply investigated by Bogacz et al. (2006), which suggested four decision rules (called *criteria for optimality* in his works) that participants may seek to maximize by modifying the speed (or the accuracy) of their response depending on internal and external factors (motivation, reward, punishment, length of the task and so on). Some of these rules are function of the subjective value for a correct and incorrect response; some others are function of the total duration of a trial; some are function of both (see **Section 2.6**). This thesis is mostly built upon Bogacz et al. seminal work, and it focuses on the investigation of decision rules.

The main aim of this thesis is to improve our understanding of decision rules. We will tackle this problem from different sides. Firstly, we will use a qualitative approach to test some predictions of the decision rules (**Chapter 4 and 5**). We will use different types of sequential sampling models, comparing their predictions when necessary. This will help us understanding what are the components affecting the decision process, across different experimental designs.

However, most of this thesis will be focussed on a new experimental framework for investigating decision mechanism and optimisation. We call this framework "EXACT Paradigm", and its main goal is to investigate decision rules without having to assume any perceptual process (**Chapter 6, 7 and 8**). We will show several applications of such design, which allows us to answer more directly the question of what decision rules are used by participants. This will help us to investigate how well participants

can estimate experimental parameters, whether they are changing their estimation during the task, and whether they are most loss-aversive, or risk-seeking. The last experimental Chapter (**Chapter 8**) provides a comparison between the EXACT paradigm and the classic approach used in decision making, by also investigating the effect of different payoff on responses.

Part of this work is connected to Piéron's Law, a fundamental psychophysics law that links the intensity of a stimulus with the speed of response (Piéron's, 1914, 1920, 1952, see **Section 2.9** and **2.10**). We will show that this rule, found in a myriad of different domains and experimental designs, is the result of the interaction between a speed-accuracy trade-off function and a decision rule. This approach predicts that, with the right experimental conditions, is it possible to break this relationship. In the later experimental Chapters we confirmed this prediction.

Another topic touched in this work regards regularities in responses distributions. We found interesting regularities both in the response rate (which led us to formulate the Rate-Hypothesis, the idea that the brain compute the response in the rate domain, and then translate it with a mechanism that resembles mechanically a sequential sampling process, see Harris et al., 2014 and **Section 2.4**) and in the accuracy distribution (which we observed within the context of the EXACT Paradigm, see **Section 7.1**). These regularities may underpin more fundamental processes going on in the brain: we will discuss the feasibility of such approach, the implication in terms of decision rules and optimality, and possible future investigations.

The field of optimality in perceptual decision making is relatively young, and only a few researchers have tackled the problem since it was first formulated by Bogacz et al. (2006). Plus, our approach diverges from the classic one as it aims to investigate optimality and decision rules by using a novel experimental design. As such, this must be considered an exploratory work, which goal is not to obtain clear answers to all our questions, but to open up different and possibly promising paths for further research.

1.1 Organization of the Thesis

Chapter 2 presents the background information that will be used throughout the thesis. Several sequential sampling models are discussed, focusing in particular on the Pure/Extended Diffusion Models. The LATER model and its connection to the Rate Hypothesis is presented. We introduce the concept of optimal decision making, discussing the work done in this area in the last decades, and present the four decision strategies used throughout this work, their similarities, differences, and the relationship that they have with the diffusion models. We then introduce the Piéron's Law, one of the main phenomena of our investigation, and its connection with optimality models.

In **Chapter 3** we present the statistical methods used in the thesis. This includes the technique for data aggregation from different participants (Vincentizing); the technique to analyse the distributions in the Rate Domain (standardising method); the mathematical description of the Pure and Extended DDM, and of the LATER model, within the Simple and Choice RT paradigm; the mathematical formulation of the four decision rules, and the respective optimum response time and accuracy by using an exponential speed-accuracy function; the technique employed to fit distributions to these models (Nelder-Mead algorithm with Quantile Maximum Likelihood) and model selection (Akaike information criterion); the technique for fitting data points, like median RT across conditions (Weighted Least Square, modified Weighted Least Square) and model selection (a modified version of the Bayesian information criterion); the shift-detection method used for analysing response signals.

Methods for analysis used for a particular experimental design, or deviation from the analysis exposed in this Chapter, are specified in the Method Section of the corresponding Chapter.

In the first experimental Chapter, **Chapter 4**, we investigate the effect of foreperiod and different contrasts on a Simple RT paradigm. We use a brightness discrimination task to obtain Piéron's Law, and analyse the behavioural result in terms of drift diffusion model and LATER model. We are particularly interested in checking what parameter in the diffusion model is affected by the foreperiod length, as this would imply different explanations in terms of decision mechanism taking place during

the task. We find that both non-decision time and threshold separation are affected by the foreperiod length, which is consistent with the idea that participant's response follows an optimisation mechanism that depends on total trial time, as some of the decision strategies assume. Alternative explanations are also discussed**.** We analyse the distribution shape in the rate domain, finding that they appear normally distributed for the long foreperiod time condition, but deviates for shorter foreperiod conditions.

In **Chapter 5** we analyse more deeply the effect of total trial time on the accuracy of response for Choice RT. We first illustrate the relationship between increasing total trial time, accuracy, response, and threshold separation. In order to separate motor preparation component due to stimulus predictability and optimization mechanism, we vary the length of one of two parts of the trial, the foreperiod part, and the response-to-warning-signal part. The study clarified the results from the previous Chapter and shows how participants were not changing their response due to the condition length, but only due to the stimulus predictability. This finding goes in contrast to several studies, and we discuss possible reasons for this. The analysis of the Rate Domain distribution is performed for both tasks, finding in most cases distributions that were overall Normal in the Rate Domain, with some caveats due to anticipations.

In **Chapter 6** we discuss the idea of conceptually dividing perceptual process and decision strategy. We suggest a solution in the form of a new experimental paradigm to analyse the decision strategy without having to rely on the perceptual process. We introduce a preliminary version of what we call "EXACT Paradigm", which hardcodes the speed-accuracy function so that it is perfectly known by the researcher (but not by the participant). We discuss the design of such paradigm, the differences with Classic RT tasks, and the experimental predictions derived from our design, focusing in particular with the prediction regarding Piéron's Law. We show the result of an experiment which investigates the viability of such paradigm and discuss how it could be improved upon.

By slightly changing the design of the previous task we introduce, in **Chapter 7**, the EXACT Paradigm, which main difference consists of making the speed-accuracy function known to the participants in the form of an increasing gauge. The results of two experiments are analysed in this

Chapter. Both tasks are designed to resemble a Piéron's Law task, by varying a parameter that mimics the stimulus strength. As a second factor, we varied the delay across trials in the first experiment and the starting point of the gauge in the second one. We found that one of the strategy (defined as RR_m in the following Chapter) appeared to be the most successful model in accounting for the observations. Furthermore, we found that, as expected, with specified experimental condition, Piéron's Law does not hold. We also suggest a new interpretation of the RT distribution that takes into account the accuracy distribution, which cannot be analysed in Classic RT task. The implications of this approach are discussed.

In **Chapter 8** we directly compare the result of an EXACT Paradigm with the result of a Classic RT paradigm, by varying the punishment/reward ratio in terms of points. Whereas in the EXACT Paradigm we found that participants were very responsive to the point feedback, the same cannot be said for the Classic RT task, where participants' average responses did not appear to be affected by it. In the EXACT Paradigm, we found a clear departure from Piéron's Law with high reward conditions, as predicted. In the Classic RT task, Piéron's Law did not hold for the low performance group, which can be easily explained by taking into account different visual acuity. In both cases, we show how Piéron's Law seem to be the result of an interaction between the speed-accuracy trade-off and a decision strategy. Finally, we perform the standard optimality analysis on the Classic RT paradigm, by using the Pure DDM as the perceptual model. In this case, the results were mostly dissatisfying.

A summary of findings is presented in **Chapter 9**. We discuss the main empirical results under the light of optimal decision making models. The possible reasons behind departure from expected decision model are discussed, including suggestions for possible future directions. A general overview of the contribution of this work to the field of decision making is provided.

Chapter 2

Background Information

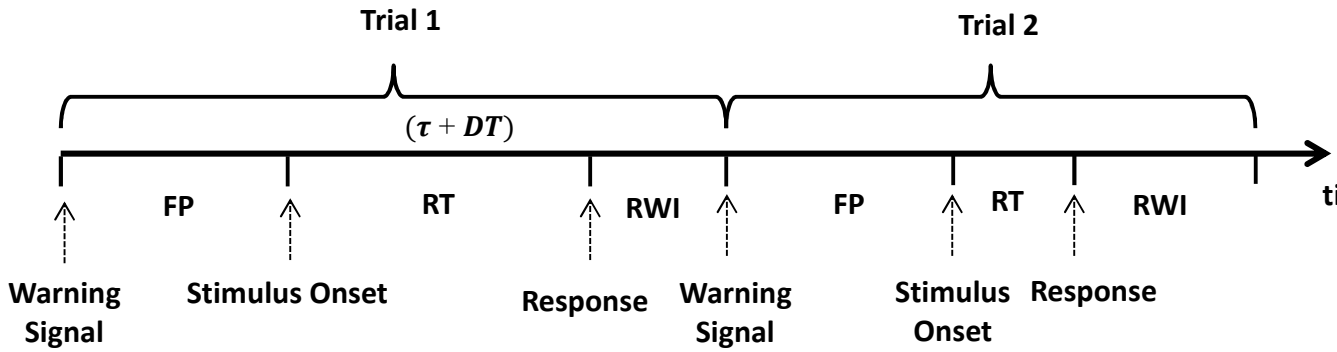
The main aim of this work is to investigate the decision rules used in the decision making processes involved in simple tasks. To do that, we are going to apply concepts related to modern models of decision making. This Chapter will provide the necessary background needed to understand the research problem and its significance. Since their first formal introduction (Bogacz et al., 2006), decision rules have been investigated through the framework of sequential sampling models, analysed by means of reaction time (RT) tasks. After a brief description of RT tasks, we will present an overview of sequential sampling models, followed by a more detailed description of the drift diffusion model, which we will extensively employ throughout this work. We will then provide the information related to the Rate-Hypothesis, another topic investigated in this work, and how the LATER model can be employed to test it. We will then describe the four decision rules tested in the present work, and their relationship with the parameters of the drift diffusion model. Finally, we describe an important phenomenon that relates reaction times with stimulus strength: the Piéron's Law. We will describe its relevance in the context of optimality, and how the analysis of decision rules can shed a new light on the underlying mechanism generating Piéron's Law.

2.1 Reaction Time Tasks

Historically, RT tasks have emerged as the standard paradigm to investigate the dynamics of decision making. Even though these tasks are extremely simplified when compared to real life situations, they are representative of many problems faced by living organisms in their natural environment. The nature of these tasks makes it possible to collect thousands of responses for a single participant in a short time, and thus allows reliable statistical analysis of rich datasets. There are many possible classifications of reaction time tasks, but the most useful for this work consists in distinguishing

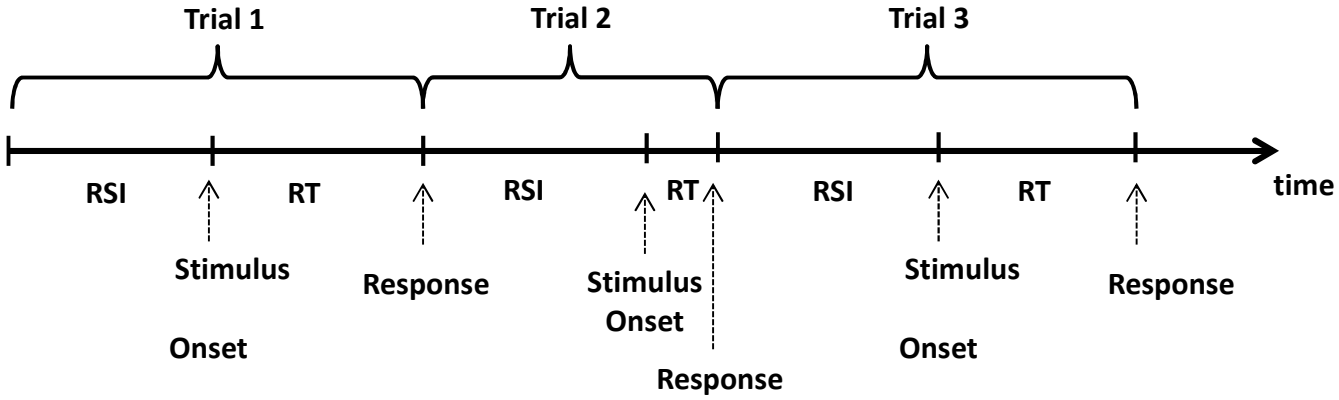
between Simple RT tasks and Choice RT tasks (Luce, 1986). In a Simple RT task (also called detection tasks) the participant is asked to press a key whenever a stimulus appear, whereas in a Choice RT task participants have to choose between two or more alternatives. When it involves only two alternatives, we will define it as a Two-Alternative Force Choice (2AFC) Task (Bogacz et al., 2006). In both cases, there are two types of designs commonly used: in the first type, the starting of each trial is indicated by a warning signal, with the stimulus (or the stimuli, if it is a Choice RT) appearing after some (constant or variable) time. The time between the warning signal and the stimulus/i is called foreperiod time (which effect will be deeply investigated in Chapter 4 and 5). After the participant's response, there is generally another pause before the starting of the next trial, called response-to-warning signal interval. This design is illustrated in the top diagram of Figure 2.1. We will employ this design in Chapter 4 and 5. The second way a RT task can be designed is by employing just one delay instead of two: in this case, after the response, the participant waits until the starting of the next stimulus onset: the participant response corresponds to the beginning of the next trial. In this case, we will refer to this delay as the response-to-stimulus interval (see bottom diagram in Figure 2.1). We will employ this design in Chapter 8.

When working with RT task, we are generally interested in the time of response and its accuracy. The time of response (the RT itself) can be divided into two components: a decision time, DT, and a sensory-motor component, τ, such that $RT = DT + \tau$. The DT corresponds to all the time in which information about the stimulus is accumulated, whereas τ correspond to the time for non-decision components, such as stimulus encoding, motor response and so on (not all models employ the τ component, e.g. in the LATER model $RT = DT$, see Section 2.4). The total time for a single trial corresponds to $DT + \tau + D$, where D corresponds to any delay in the trial that is not included in the response (for example, the response-to-stimulus interval or the foreperiod time). We also define D_{tot} as the sum of the non-decisional component, so that $D_{tot} = \tau + D$. The diagrams in the boxes in Figure 2.1 illustrates the relation between these components for the two different design.



$$D = FP + RWI$$

$$D_{tot} = \tau + D$$



$$D = RSI$$

$$D_{tot} = \tau + D$$

Figure 2.1.

Diagram of the two possible designs for a RT task. The top diagram represents the design in which a warning signal corresponds to the beginning of the trial, a foreperiod (FP) time is used from the warning signal until the stimulus onset, and a response-to-warning interval (RWI) is used after the participant's response. Note that RWI and FP may be random or constant. In the lower diagram we show an alternative design, in which only one delay is employed: the response-to-stimulus interval (RSI), from the participant's response until the next stimulus onset. In this work we are going to use both types of designs. In the box we illustrate the connection and the definition between different part of the trials for the two design. DT corresponds to the decision time, τ to the sensory-motor component, D to the delay of the trial (time that is not part of the response time), D_{tot} to the overall non-decision time.

In the past decades, several characteristics of RT responses have emerged to be particularly stable across a vast amount of tasks, designs, and sensory modalities. For instance (Brown & Heathcote, 2008; Luce, 1986; Mulder et al., 2010; Ratcliff, 2002; van Ravenzwaaij et al., 2011; Wagenmakers & Brown, 2007), mean RT is shorter for easy task than for difficult task; mean RT is proportional to standard deviation; manipulation that increases the speed of RT also increases the proportion of errors; RT distributions are right-skewed, with skew increasing with task difficulty. A typical shape of a RT distribution is shown on the left panel of Figure 2.3.

These features correspond to some law-like pattern that any model of decision making should take into account. Sequential Sampling Models have been extremely successful in explaining these patterns, as we will discuss in the next Section.

2.2 Sequential Sampling Models for Reaction Times

For decades RT tasks have been investigated using sequential sampling models (also called accumulator models, Luce, 1986). These models are unique in providing a way to understand both the speed and accuracy of responses within a common theoretical framework, whereas many models provide account for either RT (stage theory models, Townsend & Ashby, 1983; Sternberg, 1969) or accuracy (signal detection theory models, Green & Swets, 1966).

The general idea is that, during a task, the observer accumulates information based on the presented signal(s) until a stopping rule is satisfied (Luce, 1986). As the time passes, more information is gathered and therefore more accurate decisions can be made, generating a trade-off between speed and accuracy.

Specific models differ in how the accumulation of information is modelled and what is the stopping rule. These models were originally used for both Simple and Choice RT, but recently they have almost uniquely been applied to Choice RT (see Forstmann, Ratcliff, Wagenmakers, 2016). The reason is that most of these models, when applied to Simple RT, cannot take into account the anticipations (responses before the stimulus onset and, thus, a type of incorrect responses), and therefore their explanatory power is greatly reduced. Nevertheless, most of their general structure is the same for both Simple and Choice RT task. The mathematical formulation of these models for both Simple and Choice RT are presented in **Section 3.4**.

Ratcliff & Smith (2004) proposed a useful classification of these models based on the number of accumulators, and the way the information is accumulated:

- *Multiple, independent accumulators*, where the increase in the amount of evidence for each one of the N alternatives is accumulated independently by N accumulators; *single accumulator* where the information towards one alternative corresponds to information against the other alternative(s);
- *Continuous or Discrete flow* of information across time
- *Constant or Decaying* increment of accumulation of information across time
- *Stochastic or Deterministic* nature of the accumulated information.

Not all the combinations of these aspects have been implemented into a formalised mathematical model: for example, a model with decaying increment has been implemented only for continuous time and varying amount of evidence. Note also that the accumulator dependency classification loses its meaning when we use Simple RT, as a single alternative is used and thus, only one accumulator is generally assumed (e.g. Ratcliff & van Dongen, 2011).

One of the first models was developed by LaBerge (1962), for multiple choices RT task. It assumed that the evidence for each alternative is accumulated in independent "counters", and the counter that first reaches a criterion determines the response. This model was not able to correctly predict the shapes

of RT distributions, and has been consequently followed by the Accumulator Models (models in discrete time and varying amount of evidence, Smith & Vickers, 1989; Vickers, 1970, 1978, 1979; Vickers, Caudrey, & Wilson, 1971), and the Poisson Counter Models (LaBerge, 1994; Pike, 1966, 1973; Townsend & Ashby, 1983). These two models have been recently tested in comparison with diffusion models, and were not able to account for all the features of the empirical data (Ratcliff & Smith, 2004).

Another class of models based on stochastic accumulation of information was developed by Stone (1960) (see also Laming, 1968; Link & Heath, 1975; Link, 1975): the Random Walk Model, model in discrete time with varying amount of evidence and a single accumulator, which had several problems in explaining observations about RT distribution shape and accuracy (see Ratcliff & Smith, 2004). The most recent developments have been focused on continuous time models, with the most successful example being the class of Drift Diffusion Models (DDM, Ratcliff, 1978, Forstmann, Ratcliff, Wagenmakers, 2016), in which a single accumulator in continuous time accrues a varying amount of evidence with a constant drift.

2.3 Drift Diffusion Models

There are two versions of the Drift Diffusion Models (DDM) mostly used in the literature: the Pure DDM and the Extended DDM (Bogacz et al., 2006). The mathematical description of these models will be presented in Section 3.4. Here we present an informal, intuitive illustration of these models. The Pure DDM is a continuous-time version of the Random Walk Model developed by Stone (1960) and Laming (1968). As any sequential sampling model, it assumes that a stochastic accumulation of information drives the decision. This information is gathered from one accumulator (the evidence towards one alternative counts as evidence against the other alternatives). The accumulator begins from a starting point which may be biased towards one of the alternatives and rises stochastically with a certain rate (drift rate) until a threshold is reached. If the task involves a choice between one of two alternatives (2 Alternative Forced Choice Task, 2AFC) than the model includes two thresholds, one for the correct and the other for the incorrect alternative. The drift pushes the accumulator towards the correct threshold but, being noisy, it can sometime hit the wrong threshold (see top panel in Figure 2.2).

The total RT for a single trial corresponds to the sum of *decision time* plus a *residual* component, τ ($RT = DT + \tau$). In terms of diffusion model, the DT is the time from the beginning of the accumulation process until one of the two thresholds is hit. The τ component includes all the time taken for non-decisional components, e.g. stimulus encoding, memory access, motor response, which may take place before or after the decision process. We will refer to this value as "motor-sensory component". It has been found to be equal to around 100-150ms for Simple RT and 200-300ms for Choice RT (Luce, 1986). The variability in the accumulation is assumed to be normally distributed, with mean d and standard deviation s. Changing in drift rate corresponds in changing in the difficulty of the task: for example, consider a random-dot motion paradigm. In this task, a number of randomly moving dots is presented on a screen. A percentage of these dots is moving consistently to a side of the monitor, and the participant is asked to indicate the side they are moving to. In this task the amount of information gathered for each time step can be manipulated by changing the percentage of dots moving coherently. With a high proportion of moving dots (easy task), the process is assumed to have a high positive drift rate and therefore the threshold is reached quickly, resulting in high accuracy and short RT. After repeated trials, the process noise will generate a skewed distribution, like ones observed in experimental data. With a difficult task (few dots moving coherently) the process has a low drift, taking more time to reach the correct thresholds and increasing the chance to hit the wrong threshold (resulting in an incorrect response). The threshold separation (distance between two thresholds in a 2AFC, and distance between threshold and starting point for Simple RT task), denoted by a, is assumed to be under the control of the participant, and generates the speed-accuracy trade off. If the thresholds are set to be close to the starting point, responses will be faster but less accurate (the signal may hit the wrong threshold); on the other and, threshold far away from the starting point results in more accurate, but slower responses**.** The value b (from 0 to 1) defines the bias towards one alternative, where values higher than 0.5 represent a starting point closer to the correct alternative (e.g., in a random dot motion, when the direction of movement is more likely to be on one specific side). However, in our experiments there is no reason to believe that a bias is taking place, so we set $b = 0.5$ for all Choice RT experiments. The DDM model will generate response distributions that are left skewed, with the skewness increasing

with increasing difficulty, and with mean RT proportional to standard deviation RT, all features observed in real datasets (e.g. Wagenmakers & Brown, 2007; Forstmann, Ratcliff, Wagenmakers, 2016). The top panel of Figure 2.2 provides an illustration of the Pure DDM. This model has been initially used for recognition memory (Ratcliff, 1978) and has been shown to accurately predict most accuracy and RT time measures. However, an important problem is that it predicts the same distribution of correct and incorrect RT (even though in different proportions), whereas it is often seen that this is not the case (Ratcliff & Rouder, 1998; Ratcliff et al., 1999; Smith & Vickers, 1989). Until recently, the model had been set aside in favour of its extended version (discussed below). The new interest for optimum responses in terms of DDM brought it back as the standard model for RT (Bogacz et al., 2006, see **Section 2.6 and 2.7**).

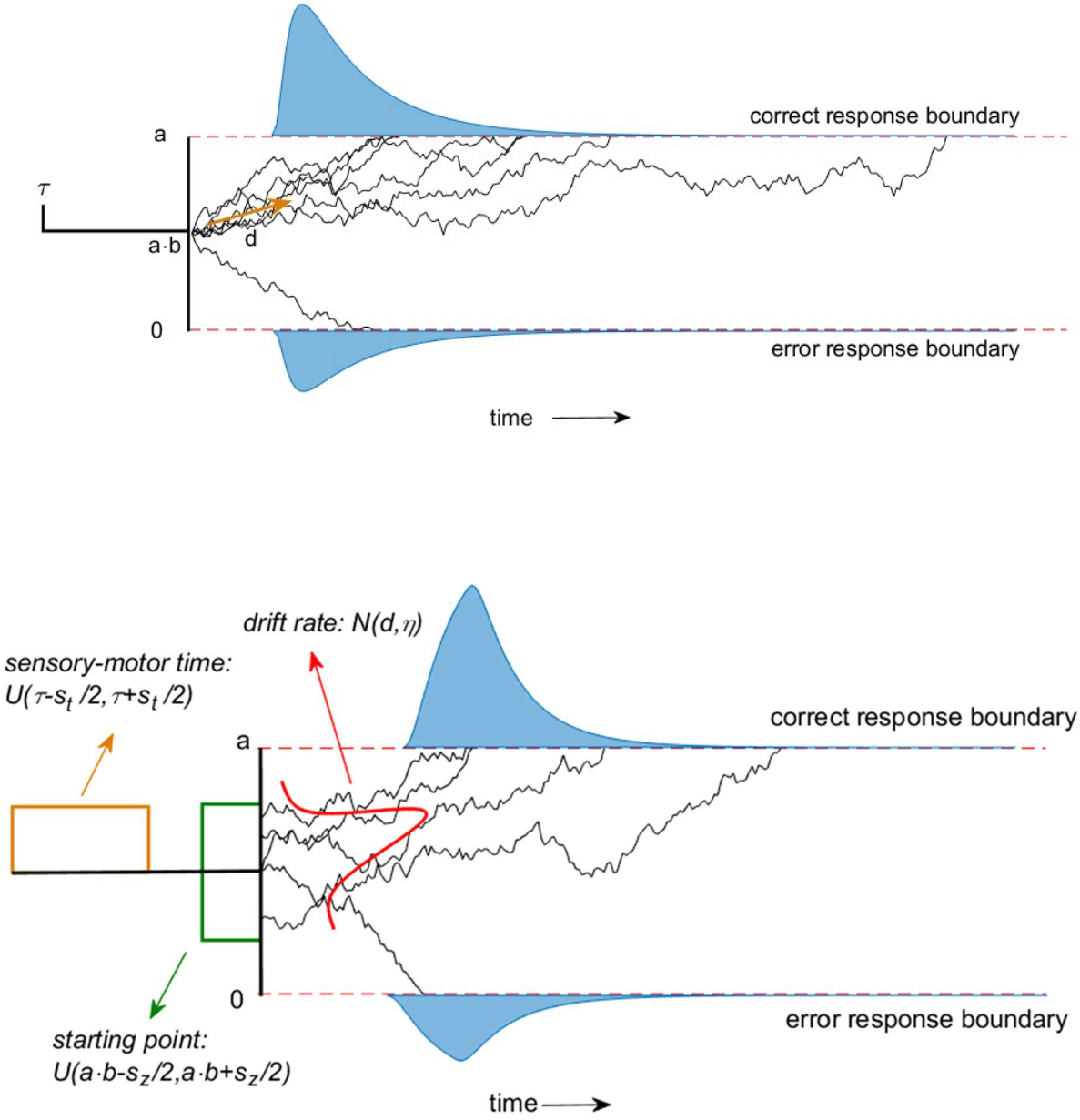


Figure 2.2

(Top) Pure Drift Diffusion Model for 2 Forced Choices Task. The process starts at a certain distance (**a** · **b**) from the threshold (**a**). The information is noisy accumulated with a drift **d** until one of the two thresholds are reached. Across trial, the drift **d** is kept constant. The resulting RT distributions (in blue) are the same for both correct and incorrect response. A constant non-decision time **τ** is added to the process. **(Bottom)** Extended Drift Diffusion Model for 2 Forced Choices Task. The process start at a point drawn from an uniform distribution (green rectangle) with a drift changing from trial to trial (and drawn from a Normal Distribution, in red). A random sensory-motor time (drawn from another uniform distribution, in orange), is added. This results in different distribution for correct and incorrect responses (in blue).

An accurate account of the relationship between response time for correct and incorrect decisions can be obtained by increasing the complexity of this model, obtaining the Extended Diffusion Model, illustrated in the bottom panel of Figure 2.2. Three sources of variability across trials are added: the starting point, the drift rate and the residual time are now random variables which value fluctuates from a trial to another. The drift rate is usually assumed to vary across trial according to a Normal Distribution with mean ξ and standard deviation η, and generates incorrect responses that are *slower* than correct responses. The variability across trial in the starting point is uniformly distributed with mean z and range s_z, and allows the model to predict error responses that are *faster* than correct responses, but only on high levels of accuracy. These predictions match what found in empirical observations (Ratcliff & Rouder, 1998; Ratcliff et al., 1999; Smith & Vickers, 1988). Another strength is that this model does not generate errors that are slower than correct response in high-accuracy conditions and faster in low-accuracy conditions, a pattern that has never been observed in RT tasks (Ratcliff & Smith, 2004). Finally, the non-decision time is assumed to vary according to a uniform distribution with mean τ and range s_t, and it allows the model to generate distributions with different leading edge (steeper or smoother), as shown in Ratcliff, Gomez, & McKoon (2004). The extended DDM can account for an impressive amount of behavioural and neuroscientific data including lexical decision making (Ratcliff, Gomez, & McKoon, 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008), recognition memory (Ratcliff, Thapar, & McKoon, 2004), effect of practice (Dutilh, Vendekerckhove, Tuerlinckx, & Wagenmakers, 2009), reading impairment (Ratcliff, Perea, Colangelo, & Buchanan, 2004), simple perceptual tasks (Ratcliff & Rouder, 1998; Ratcliff, 2002; Brown & Heathcote, 2008; Usher & McClelland, 2001). Most recently, the DDM has been analysed under the framework of neural firing patterns, suggesting a linking between the two. For example, Shadlen & Newsome (2001) found, in monkeys performing a saccadic 2AFC task, that the time at which the firing rate in the lateral intraparietal area (which is related to gaze direction) reaches a fixed threshold, predicted the response (see also Ratcliff, Hasegawa, Hasegawa, Smith, & Segraves, 2007; Gold & Shadlen, 2002; Hanes & Schall, 1996; Schall & Thompson, 1999; Shadlen & Newsome, 2001).

Regardless of the amount of research performed with the DDM, all the aforementioned studies were performed by using a Choice RT design. Applying Pure or Extended DDM to Simple RT means simply to assume one boundary instead of two. For the Extended DDM, this implies that a percentage of signals will never reach the threshold (as the drift may sometime be negative). To solve this problem, either signals with negative drifts are ignored altogether, or a simulated "signal deadline" may be used. Only few studies have used the DDM for Simple RT task. Ratcliff & van Dongen (2011) used the Extended DDM (with one boundary) to account for changes in RT to brightness level and sleep deprivation in a Simple RT task. A similar analysis was performed by Patanaik et al. (2014). The one-boundary version of the Extended DDM was also used to model a simple driving task (Ratcliff & Strayer, 2014). In these works, it has been noticed that the variability in the starting point (s_z) does not appear to improve the goodness of fit, and is thus ignored (Ratcliff & van Dongen, 2011). As far as we know, the Extended DDM for Simple RT has been usually applied without any comparison with the simpler Pure DDM. This is peculiar, as the main reason for the Extended DDM to be adopted in the first place was to explain the difference between correct and incorrect distribution. In Simple RT there is no "incorrect" distribution, and thus the Pure DDM may be perfectly fit to account for the data.

We finally note a general limitation of these models for Simple RT: neither the Pure nor the Extended DDM can account for anticipations. Recall that, for both Simple and Choice RT, and for both Extended and Pure DDM, the starting point of the sampling process corresponds to the stimulus onset. Whereas this seems reasonable for Choice RT, as the time preceding the appearance of the stimulus does not appear to give any relevant information to participants, it seems more arbitrary for Simple RT, as participants are probably basing their choice on the flow of time itself (as we will see in the experiment in Chapter 4). By starting the process at stimulus onset, the model cannot account for anticipation. We briefly discuss a solution for this issue in **Section 4.4**.

2.4 Rate-Hypothesis and LATER Model

It has been observed, within the context of saccadic eye movements, that the speed of saccades has frequency distributions that are remarkably close to the reciprocal Normal (Carpenter, 1981; Carpenter

& Williams, 1995; Reddi & Carpenter, 2000). That is, the reciprocal of saccade latencies (1/latency) has a near-Normal distribution (see Figure 2.3)**.** This led Harris et al. (2014) to hypothesise that the brain is estimating a preferred rate of response for a given set of experimental condition, and the accumulation part of the aforementioned sequential sampling models is nothing else but a mechanism to convert rate in time. From this point of view, the skewed distribution in the time domain is an artefact of the decision process, whereas the generation of a Normal Distribution in the rate domain is the main goal of the decision process itself. This is a tempting hypothesis, especially considering that one of the decision strategies investigated in the literature (and that we are going to apply to our experiment) maximises the reward rate (**Section 2.6**).

Several saccadic studies found that distributions are not perfectly Normal, but they usually present some very short responses that do not follow the main latency distribution (Carpenter & Williams, 1995; Noorani, 2014). These responses have been named 'express latencies' (Carpenter & Williams, 1995; Fisher & Boch, 1983; Fischer & Ramsperger, 1984), and they may possibly come from a different decision mechanism, for example one based on time estimation and not on perceptual process. Similarly, fast responses have been observed for manual RT distributions as well (Ratcliff, 1993; Ratcliff & Tuerlinickx, 2002; Ulrich & Miller, 1994). These responses have also been called "non-integrative responses", as they may be not the result of the sequential accumulation of information but of some other process involved in the decision making mechanism (Simen et al., 2009). As discussed in **Section 3.3**, we used several strategies to take these extraneous responses into account when analysing the distributions in the rate domain.

All the previous mentioned studies that found near-normality distribution in the rate domain have been based on saccades latencies. We have recently verified (Harris et al., 2014) that similar results hold for Manual Choice RT. In particular, by excluding severely truncated distributions we found a near-perfect Normal distribution in the rate domain, with a slight deviation in the tail (around the lower 10% of responses), which is similar to the early responses found with saccades and deemed to be caused by the "express" saccades aforementioned. However, investigation of rate domain has been generally

neglected for manual RT, and to the extent of our knowledge no study has been performed with Simple RT. In this work, we will analyse the rate distributions for both Simple and Choice RT, across different experimental conditions, in this way trying to find evidence for or against the Rate-Hypothesis. We will use two methods for doing this: aggregating participants together and computing the rate distributions will allow us to check whether distributions appear Normal in the rate domain (we will use the Kolmogorov-Smirnov test for Normality, see **Section 3.3**). The second method involves fitting the LATER model to the data.

The LATER model (Linear Approach to Threshold with Ergodic Rate) is a sequential sampling model in which a deterministic, constant accumulation of information drives the signal. More formally, at the stimulus onset, a decision signal S start rising linearly at rate r from an initial level S_0 until a threshold S_T is reached. By assuming that the rate r fluctuates from trial to trial according to a Normal Distribution with mean μ and standard deviation σ, we obtain RT distributed as $\frac{S_T - S_0}{\Phi(\mu,\sigma)}$, where Φ is the Normal probability density function. Within this model, the 'express saccades' would be explained as coming from another LATER decision process with a mean rate of 0 and very large σ (Noorani & Carpenter, 2011; Noorani, 2014). For mathematical details, see **Section 3.5**.

Similarly to DDM, the LATER model can explain how latencies are affected by experimental instructions (e.g. speed vs accuracy changes the S_t level, Reddi & Carpenter, 2000), or by stimulus strength (by changing the mean rate of accumulation μ, Carpenter, 2004). In contrast to the DDM model, the LATER model does not include a motor-sensory response added to the main decision time, as this would not result in a normal distribution in the rate domain.

The LATER model is particularly suited for testing the Rate-Hypothesis, as it is a simpler version of the DDM that, mathematically, is bound to generate Normal Distribution in the rate domain. We will use this model to compare it with the results from the DDM. The LATER model has obviously less flexibility than the other models, so that it will, in general, provide a worst fit than the DDM. However, our model selection method (AIC, see **Section 3.7**) should tell us if this is balanced out by its simplicity.

This should provide some insight on whether the underlying decision process computes the response in time or rate domain. However, as we will see in **Section 5.4.4**, distinguishing between a Normal Distribution and a rate distribution generated by a DDM (with the right set of parameters) may be extremely difficult. Thus, observing a rate distribution is a necessary but not sufficient condition to support the Rate Hypothesis.

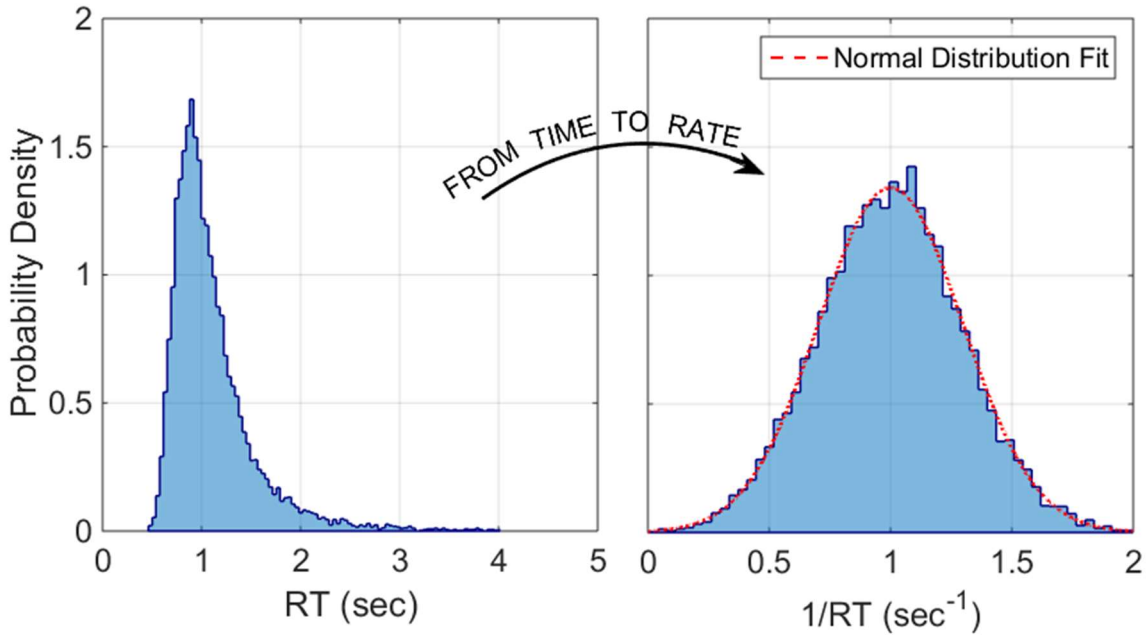


Figure 2.3

Simulated Reaction Time distribution in time domain **(left panel)** with their peculiar left-skewed shape. It has been observed that by converting the response we obtain an approximately Normal Distribution **(right panel)**.

2.5 Other Sequential Sampling Models

There are still three models that deserve to be mentioned: the Ornstein-Uhlenbeck (O-U) , the Linear Ballistic Accumulator (LBA, Brown & Heathcote, 2008) and the EZ-diffusion model (Wagenmakers, van derMass, & Grasman, 2007). The O-U model adds to the DDM mathematical formulation a third term, linearly dependent on the value of the signal at time t, which allows the prediction of "risky" or "conservative" behaviour depending on the value of this added parameter. The Linear Ballistic Accumulator assumes multiple and linear accumulators which accumulate evidence for two or more responses. The starting point and average speed of each accumulator vary from trial to trial, similarly

to the Extended DDM, but the average speed is bounded to be positive, as only one threshold is present. A comparative study between LBA and DDM is presented in Donkin et al. (2011). The EZ-diffusion model is a model with a unique linear accumulator with two thresholds, and within-trial variability in the rate of accumulation. This model lacks the power to account for important empirical phenomena in the data (Brown & Heathcote, 2008). We decided to limit our analysis to DDM and LATER for these reasons:

1) As seen above, the Extended DDM has been applied to a variety of task and sensory domain, and it always provided an excellent fitting to the real data. Its ability to fit almost all types of features of RT distributions is unmatched by other models.

2) Even though the Pure DDM cannot take into account certain features of the response distributions, analytical solution for optimal responses for several decision strategies can be easily derived, which makes it an excellent model for investigating optimality in decision making.

3) In this thesis, the LATER model is used not as an alternative sequential sampling model, but as a way to provide an additional test of the Rate Hypothesis. By fitting the LATER model, we directly compare its goodness of fit with the other two DDM models.

2.6 Optimal Decision Making: Decision Strategies and Speed-Accuracy Function

An important part of any sequential sampling model is the stopping rule (which corresponds to the decision thresholds in models such as LBA or DDM). Informally, the stopping rule can be seen as a way to implement the strategy that the participant uses in order to trade speed for accuracy. For example, in a difficult task, the participant may either decide to accumulate a large amount of information before responding (but spending a lot of time doing so) or, in the extreme case, responding at random, as fast as possible, in order to have more trials. By modifying the threshold, the participant can implement different stopping rules. But *when* to use one rule or the other depends on the exact form of the strategy. For example, it seems reasonable that participants should increase the threshold with higher punishment, in order to be more accurate. But by how much, exactly? And is this strategy always valid, or only for

a certain range of punishment values? Some other cases may be even less intuitive. For example, should the participants increase the threshold (accumulate more information and thus taking longer to respond) with longer trial length? The first systematic approach to the decision strategies was presented by Bogacz et al. (2006). In that theoretical work, four decision strategies (called "criteria for optimality") were proposed, and have since then tested in the related works (Balci et al., 2011; Zackenhouse, Bogacz, & Holmes, 2010; Bogacz et al., 2010; Simen et al.; 2009). These decision strategies, which we will discuss and test extensively in this work, are:

Reward Rate (RR): defined as the proportion of correct trial divided by the average duration between trials (Bogacz et al., 2006). Responses do not depend on the reward/punishment policy used in the task.

Modified Reward Rate (RR_m): proposed by Bogacz et al. (2006) and Harris et al. (2014), is it defined as the proportion of total utility divided by the average duration between trials. Differently from RR, this decision rule takes into account not only the proportion of correct responses, but the incorrect ones as well, and scales them by the corresponding reward and punishment (in terms of subjective utility). Harris et al (2014) have hypothesized that the RR_m may account for the observed normality distribution in the rate domain (Rate-Hypothesis, see **Section 2.4**). RR_m is also an optimal strategy for tasks with fixed condition length.

Reward/Accuracy (RA): the formal version of the COBRA theory of Maddox & Bohil (1998), formalised by Bogacz et al. (2006). It is a weighted differently between RR and accuracy, and it is also function of total length of the trial. In the original version by Maddox and Bohil (1998) it is assumed that participants attempt to maximize reward in the long-run, but places some importance on the response accuracy as well. Qualitatively, this decision rule is similar to RR_m, as they are both function of the total delay and depend on the emphasis on accuracy; quantitatively, RA generates different predictions, and in general is not an optimal strategy for tasks with fixed condition length, like RR_m.

Bayes Risk (BR): initially theorised by Wald and Wolfowitz (1948) in order to prove optimality for the Sequential Probability Ratio test, it assumes that decision makers seek to minimise the weighted sum of time and error rate, but without taking into account the total length of the trial;

These decision rules describe what is the "quantity" that the participant is trying to maximise. Each decision rule is a function of time and accuracy, so that given a specific relationship between accuracy and time there is a unique time of responding which maximises it the corresponding decision rule. We call this optimum decision time DT^*. In order to respond optimally, participant should respond as close as possible to the maximization point. The relationship between speed and accuracy, the speed-accuracy trade-off function, is defined as $ACC(DT)$. Intuitively, the rate of increase in accuracy depends on the stimulus strength (in a 2AFC, on the difference between the two alternatives), so that higher stimulus strength corresponds to steeper increase. Depending on the perceptual mechanism assumed, $ACC(DT)$ may also depend on the stimulus noise, the number of alternatives, or other experimental characteristics.

Decision rules can be also function of two more parameters: the emphasis on accuracy q and the total non-decision time D_{tot} . Greater values of q correspond to greater emphasis on accuracy (Bogacz et al., 2006, but note that RR is not function of q, as only correct responses are taken into account), but the precise mathematical formulation is different with different decision strategies: with BR and RA, it describes the weighting between accuracy relative to speed, whereas in RR_m it describes the subjective punishment to reward ratio, that is how much weight does the participant give to an incorrect response compared to a correct one. This can be affected by monetary rewards or verbal instruction, as we will see in Chapter 8. Increasing this parameter corresponds to an increase in the optimum response time. However, when coupled with different stimulus strengths, varying the q parameter may generate the interesting phenomenon of "breaking" Piéron's Law (see next Paragraph, **Section 2.9**, and **Section 2.10**). Another parameter shared by RR, RR_m and RA (but not BR) is D_{tot}, which corresponds to the time of the trial which is not decision time, that is, all the time that does not contribute to an increase in accuracy. We define $D_{tot} = D + \tau$, where D is any delay across trials (for example, the time between a

response and the next trial), and τ is the sensory-motor component. Like q, increasing D_{tot} corresponds to an increase in the optimum response.

Figure 2.4 provides a summary of the conceptual blocks needed to go from the speed-accuracy function to a DT^* curve (defined as a series of optimum decision times given several values of stimulus strength). On the top row a simple speed-accuracy function $ACC(DT)$ is shown. In this case we used an exponential function, similarly to the one used in Chapter 6, 7 and 8 (**Equation 6.1**). By increasing stimulus strength, the accuracy will increase faster in time. The middle rows panels show the value of each decision rule given time, and for different stimulus strength (by keeping the other parameters q and D_{tot} fixed). The red filled dot indicates the maximum value for that decision rule, at which time the participant should respond to maximize it. Note how the optimum decision time is 0 for very low stimulus strength, and then increases with increasing stimulus strength before decreasing again. The intuitive explanation is that with low stimulus strength there is no much information to be accumulated, so waiting would only waste (precious) time that could be used to perform more trial. With increasing stimulus strength, it starts to be convenient to accumulate information, until the stimulus strength become so high, and the accuracy increases so fast, that responses can decrease again. This relationship is consistent across RR, BR and RA, but it holds for RR_m only for some combination of parameters, which makes this decision rule particularly interesting (see **Section 2.10**; this will be discussed in greater details in **Section 6.1.4** and **6.1.5**). The blue lines in the middle row panels connect the optimum decision time for several values of stimulus strength and a fixed set of experimental condition, and correspond to a single DT^* curve. In the bottom row panel, we show a set of DT^* curves for the RR_m decision rule. The blue line shown here corresponds to the same DT^* curve in the modified Reward Rate panel in the middle row. The other curves correspond to DT^* curves generated by different experimental conditions (in this case, by increasing the punishment/reward ratio, q). A similar DT^* curve plot, for each one of the other decision rules, is shown in Figure 6.2.

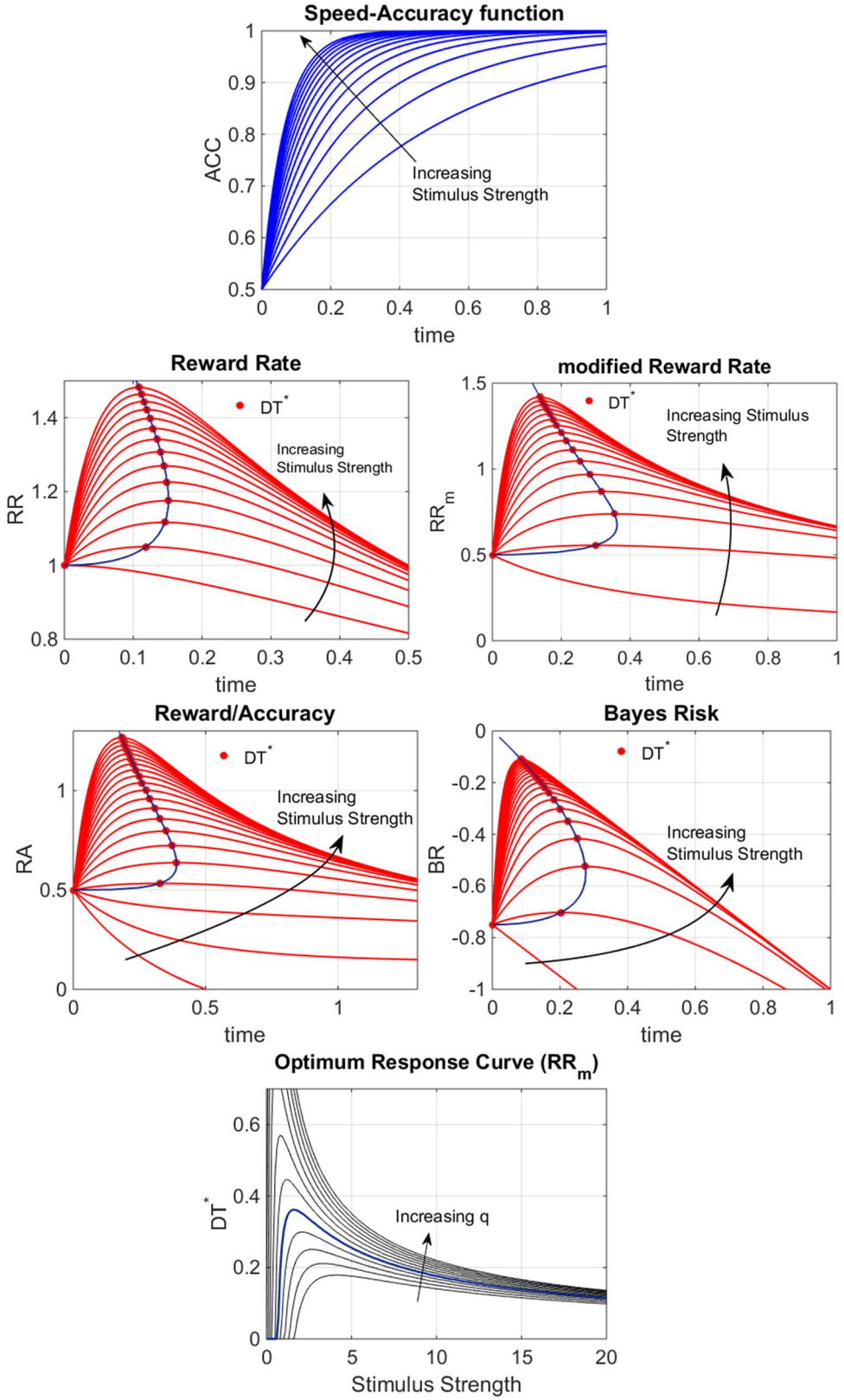

Speed-Accuracy function
ACC
time
Increasing
Stimulus Strength
Reward Rate
RR
time
DT*
Increasing
Stimulus Strength
modified Reward Rate
RR_m
time
DT*
Increasing Stimulus
Strength
Reward/Accuracy
RA
time
DT*
Increasing
Stimulus Strength
Bayes Risk
BR
time
DT*
Increasing
Stimulus Strength
Optimum Response Curve (RR_m)
DT*
Stimulus Strength
Increasing q

Figure 2.4

Speed-Accuracy function **(Top Row)** describe the relationship between speed and accuracy. As an illustrative example, we used an exponential function. The accuracy grows from 0.5 (50% of probability of being correct at chance level, with a 2 alternative task) to 1. By increasing the Stimulus Strength we obtain a function that increases faster with time. Decision Rules' value **(Middle Rows)** show the value of each decision rule in time, for a set of stimulus strength, based on the speed-accuracy function above. The red filled circles indicate the maximum value for each stimulus strength condition. This generates an optimum decision time (DT^*) curve **(Bottom Row)**, in blue in the middle rows and in the bottom row. In the bottom row we also show how manipulating the emphasis on accuracy parameter affect the shape of the DT^* curve. RR_m = modified Reward Rate.

In order to provide a quantitative analysis of the decision strategy, a precise formulation of the relationship between speed and accuracy ($ACC(DT)$) needs to be defined. $ACC(DT)$ is usually believed to be a concave, monotonically increasing function, reaching 1 asymptotically, similar to the one we used in Figure 2.3 for the purpose of illustration (Luce, 1986). The usual approach is to use the relationship defined by the DDM (see **Section 8.5.1**), and base the decision rule predictions upon it. To the extent of our knowledge, only the Pure and Extended DDM have been used for this purpose. However, decision rules and DDM are two independent concepts, and, in principle, decision rules can be analysed using any $ACC(DT)$. The problem in assuming a particular sequential sampling model is that, as we have seen in the previous sections, there are a variety of possible perceptual mechanisms that have different mechanism and thus different $ACC(DT)$. With different perceptual mechanisms, the same decision rule generates different predictions. Therefore, when empirical observations and predictions do not match, is it not possible to say if it depends on the decision rule itself or on the perceptual mechanism assumed. Comparing several decision processes and sampling models can be extremely difficult and, even in this case, it is not possible to know whether the perceptual mechanisms used comprises the one employed by human participants. In Chapter 6, 7 and 8 we tackle this problem by developing a new paradigm in which decision strategies are analysed across several participants without having to infer any perceptual mechanism, but by "hard-coding" the perceptual mechanism within the experiment. This paradigm is referred as to the EXACT (EXogenous ACcumulation Task). In Chapter 8 we try to match experimental findings of the EXACT paradigm with a classic RT design, by fitting different decision strategies to the Pure and Extended DDM. In the method section of Chapter

8 we will also provide more mathematical details relative to the connection between the DDM and the decision strategies.

2.7 Analysing the Decision Rules: Three Levels of Analysis

A formal description of the mathematical methods for analysing the decision rules will be presented in the next Chapter and, for the DDM, in **Section 8.1**. Here we want to illustrate three different ways in which is possible to make inference about the decision rules used. At the first level, we have the qualitative description of the effect of decision rules on RT and accuracy by changing experimental conditions. For example, we know that by affecting the D_{tot} (for example, by increasing the waiting time across trials), the decision rules RR, RR_m and RA predict that both RT and accuracy increases. This is a very general description that does not require using any particular perceptual process, but it requires only some (very) general assumptions about the shape of the speed-accuracy function (in this case, that is monotonically increasing). In this way we can distinguish between different groups of decision rules (for example, those affected by D_{tot}, and those who are not).

To get to the second level of description, we need to have a precise perceptual model process, which gives a relationship between RT, ACC and the process' parameters. In this case we can predict that, for example, by increasing D_{tot} we are going to observe an increase in RT, ACC, and a consequent increase in the threshold separation. This is the approach we are using in Chapter 4 and 5 (see **Section 2.8** for a general description of the relation between parameter of the DDM and decision rules). The analysis will still be qualitative.

To be able to obtain quantitative predictions we need an experimental design in which each condition generates distinctive error rates (third level of analysis). In this case we can not only estimate perceptual process' parameters, but we can compare them with the theoretical optimum parameter with each decision rule. We are not only able to distinguish between different groups of decision strategies, but also between each decision strategy. This approach requires several assumptions, and entails fitting the perceptual model assumed, a procedure that can be complex and may introduce noise. We are going

to apply this method for the last experiment in Chapter 8. The risks of this approach are discussed in Chapter 6 and, from that point on, we will mostly employ a different methodology that allows a precise, quantitative estimation, without having to assume a particular decision process (the EXACT Paradigm).

2.8 Drift Diffusion Model and Optimality

One of the main goals of the present work is to find a way to properly separate the decision rule with the underlying perceptual process. We achieve this by using the EXACT Paradigm in Chapter 6, 7 and 8. However, it is important to consider the implication of assuming the drift diffusion model as the perceptual model used by participants, as this is still a useful tool when investigating Classic RT observations. In Chapter 8 we will fit the results to the aforementioned decision strategies, and the mathematical details will be presented there. For the experiments in Chapter 4 and 5, however, the number of errors is not large enough for a quantitative analysis, and we have to resort to qualitative predictions, by relating the increase of DDM parameters to different experimental conditions.

The basic DDM assumption is that the bias parameter (b), drift rate (d) and sensory-motor time (τ) are affected by the experimental design and are not under the direct control of the participant, whereas the threshold (a) can be consciously changed according to experimental instruction/subjective utility payoff/change of strategy or other external or internal variables (Ratcliff & Rouder, 1998). Recall that by increasing the threshold, the DDM generates responses that are slower but more accurate. We are now going to illustrate how participants can change the threshold for different sets of parameters in order to maximize the value described by the decision rule used. The first column of Figure 2.5 shows the non-monotonic relationship between threshold and drift rate. With low drift rate, there is no information available to be accumulated, and so it is convenient to guess immediately. As the drift rate increases, it becomes advantageous to accumulate information, and the threshold increases. With high enough drift rate, the signal dominates the noise, so that the threshold is lowered again until the decision can be made almost immediately (Bogacz et al., 2006). This has a similar effect to the relationship between stimulus strength and RT illustrated in the bottom row of Figure 2.4. The second column of Figure 2.5 shows a monotonic dependence on non-decision time D_{tot} (the total delay, including the

sensory-motor component, $D_{tot} = D + \tau$), so that as D_{tot} decreases towards zero, the optimal threshold also decreases toward zero (assuming that the parameter q is less than one, so that the subjective reward is higher than the subjective punishment). As D_{tot} increases, the optimal threshold value increases indefinitely. Similarly, for greatest emphasis on accuracy (q), the optimal threshold grows indefinitely, as shown in the third column of Figure 2.5. In Chapter 4 and 5 we will vary the total trial time, therefore changing the parameter D_{tot}, and check whether the estimated threshold parameter changes accordingly. In Chapter 8 we will instead try to affect the emphasis on accuracy (q) and, similarly, check whether this corresponds to an increase in the estimated threshold.

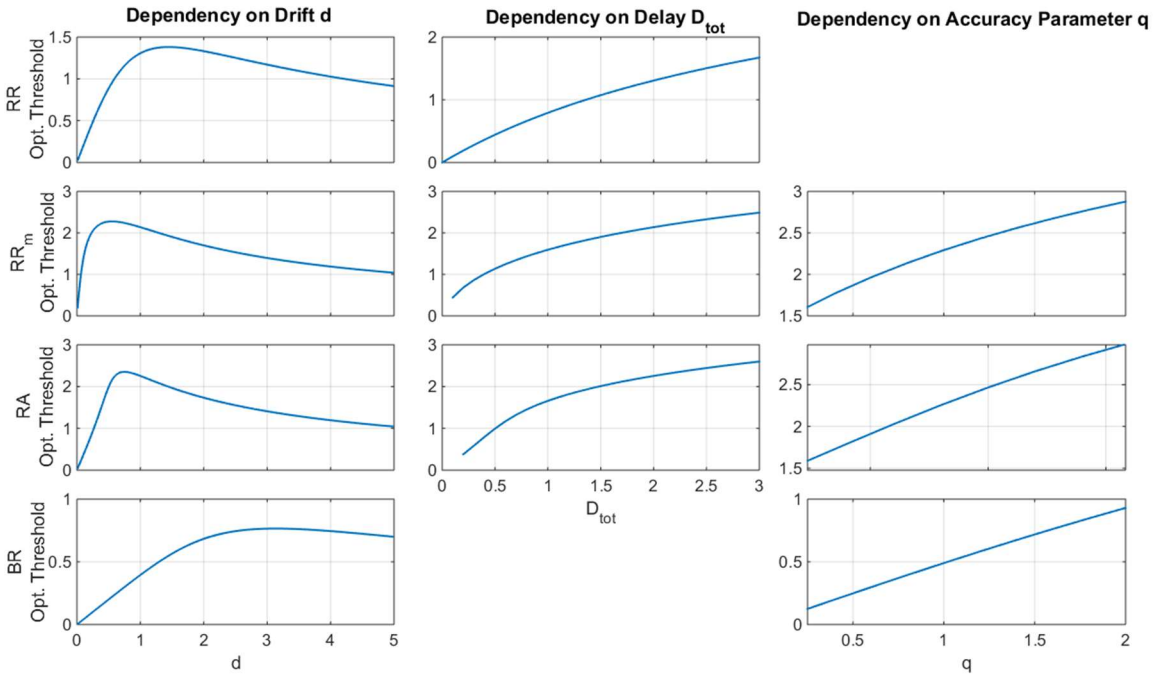


Figure 2.5

Optimum threshold for each decision rules (on the vertical axis) and different values of the Pure DDM parameters (indicated on top of each column), with the other parameters held fixed at $\boldsymbol{D = 1, D_{tot} = 2s, q = 0.8}$. Not that the relationship is monotonically increasing for the non-decision time $\boldsymbol{D_{tot}}$ and the emphasis on accuracy parameter $\boldsymbol{q}$, but non-monotonic for the drift rate $\boldsymbol{d}$. $\boldsymbol{RR}$ = Reward Rate; $\boldsymbol{RR_m}$ = modified Reward Rate; $\boldsymbol{RA}$ = Reward/Accuracy; $\boldsymbol{BR}$ = Bayes Risk; Opt. = Optimum.

2.9 Piéron's Law

Piéron (1914; 1920; 1952) first described quantitatively the relationship between RT and intensity of the visual stimulus. This relationship took the name of Piéron's Law and states that the mean RT decreases as a power law with increasing stimulus intensity I.

$$mRT = kI^{-\beta} + \gamma$$

Where k and β are scaling parameters that determine the slope of the function and γ is the intercept (Luce, 1986). Figure 2.6 presents some instances of Piéron's Law with different λ and β. During the last century, the law has been reported in many different domains, including brightness detection (Pieron, 1914), tone detection (Chocholle, 1940) taste detection of dissolved substances (Bonnet et al., 1999), odour detection (Overbosch et al., 1989), heat detection (Banks, 1973), colour saturation (Stafford et al., 2011) simple and choice reaction times (Pins & Bonnet, 1996; 1997; 2000), go/no go tasks (Jaskowski & Sobieralska, 2004), Stroop conflict task (Stafford et al., 2011), two-choice dot-motion perception task (Maanen et al., 2012), random-dot kinematogram (Reddi, Asrress, & Carpenter, 2003). This law is so general that the stimulus intensity parameter should not be interpreted as the luminous intensity of a wavelength (as initially assumed by Piéron), by as a general stimulus strength, that is any stimulus feature that affects how fast the accuracy is accumulated.

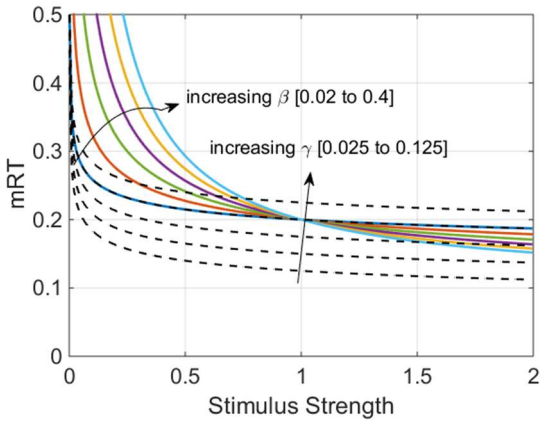


Figure 2.6

Different examples of Piéron's Law. Changing the parameter γ corresponds in shifting the function up and down (dashed black lines), whereas changing the β parameter corresponds to changing the slope of the function (coloured lines). mRT = mean Reaction Time.

There are three explanations to account for Piéron's Law. Historically, this effect has been considered related to a physiological stimulus scaling (Nachmias & Kocher, 1970), but the ubiquity of so many paradigms and sensory modality has lead researcher to consider alternative explanations. For example, an information-theoretic approach derives Piéron's equation from an optimal information process in sensory perception (Medina, 2009). Stafford and Gurney (2004) firstly proposed that Piéron's Law may arise by an accumulator models. This hypothesis has been tested with the Linear Ballistic Accumulator by Donkin and Maanen (2014). The author used a "free" model (where the drift rate was freely estimated for each luminance level) and a linear model (linear relationship between drift rate and luminance level). The results suggested that the drift rate was not linearly related with the proportion of white pixels (for a brightness discrimination task) or angular distance (for a random-dot motion task). Similarly, Carpenter, Reddi and Anderson (2009) found that rates of accumulation in the LATER model were not linearly related to contrast levels. The overall conclusion is that Piéron's Law was not simply the result of the rise-to-threshold architecture, but reflects a non-linear relationship between stimulus strength and stimulus representation (drift rate, in the accumulator model). However, we do not agree with this interpretation. Finding that some stimulus characteristic does not have a linear relationship with drift rate does not say, by itself, very much, as this would depend on the characteristic chosen. For example, Donkin et al. (2014) chose, for their brightness detection experiment, the percentage of white pixels. However, they could have chosen luminance, or contrast, or any other stimulus value describing its strength. It would have indeed been surprising if the percentage of white pixels, among the myriad of stimulus property, had a linear relationship with drift rate. The main point is that a rise-to-threshold model can, in fact, generate a Piéron's shaped function with most monotonic stimulus-drift rate relationship. This does not imply that Piéron's Law needs to reflect a different mechanism, or that it needs a different explanation. It implies that there is some unknown relation between the stimulus

characteristic we chose and the drift rate. We believe that, while Piéron's Law can be in fact generated by a sequential sampling model, it is deeply connected with the decision rule used.

2.10 Piéron's Law and Optimal Response

We noted that by applying one of the four decision rules to any monotonically increasing and concave speed accuracy function $ACC(DT)$ we can actually obtain a series of optimum responses that matches Piéron's Law. An illustration of this is shown in Figure 2.7, in which we used RR_m as a decision rule and an exponential ACC (Equation 6.1). Note that Piéron's Law matches the DT^* curves for any stimulus strength only given some condition of the experimental design for RR_m (left panel of Figure 2.6, the top curve can be fitted by a Piéron's Law from the lowest values of the stimulus strength, to the highest). This condition regards the starting point of the $ACC(D$ function and the subjective punishment/reward parameter q (this will be more formally discussed in **Section 6.1.4** and **6.1.5**). When these conditions are not met for RR_m, and *in any case* for the other decision rules, Piéron's Law matches the DT^* curves only from a certain point on (after the maximum point) In Figure 2.6, we fitted two DT^* curves of RR_m and all the DT^* curves from RR starting from an arbitrary point slightly over the maximum point of the curve. Piéron's Law fit similarly well for BR and RA decision rules.

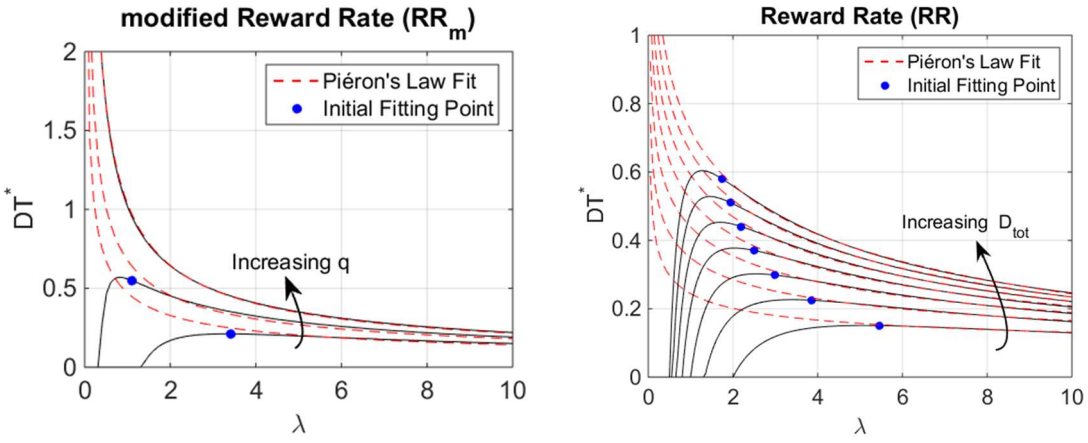


Figure 2.7

Fitting Piéron's Law to optimum decision time (DT^*) curves for. The optimum decision time curves (black lines) for the modified Reward Rate **(left panel)** and the Reward Rate **(right panel)** are shown. Note that for the modified Reward Rate we can obtain a DT^* curve that actually goes to infinity with low stimulus strength, as predicted by Piéron's Law, and thus obtaining a perfect fit. For the other conditions, we can still obtain a reliable fit if we consider the DT^* starting from a high

enough stimulus strength (in the panel, indicated by the blue circles). q = emphasis on accuracy parameter; D_{tot} = non-decision time = $D + \tau$.

In this work, we are going to see how Piéron's Law results from applying one of the aforementioned decision strategies to a monotonically increasing speed-accuracy function. Intuitively, stronger stimuli will make accuracy increase faster, so that a certain level of accuracy can be reached faster (shortest RT with higher stimulus intensity). Piéron's Law predicts that, with lower and lower stimulus intensity, RT will increase to infinity. On the other side, all the four decision strategies that we have mentioned predict that when the stimulus strength is too low, the best strategy is to respond faster (we stress that, with RR_m, this is true only for a certain combination of parameters). This generates a non-monotonic relationship between RT and stimulus strength, which we will refer to as "non-Piéron's shape". The interesting point is that in all the myriads of Piéron's-type experiments, this non-monotonic relationship has never been observed. Observing it would confirm that Piéron's Law is not a pure physiologic mechanism, but is mediated by the decision strategy employed. Addressing this phenomenon will be one of the main focus of Chapter 7 and 8.

Chapter 3

Statistical Analysis and Model Fitting

In this section, we illustrate the methods used across most of the next chapters for analysing behavioural data and to perform model fitting/model selection. We will describe the mathematical formulation of the Pure/Extended DDM and the LATER model, and explain the routine used for fitting and for selecting the best model. The procedure for fitting a set of observations is also explained. Models or models' details used uniquely in one chapter are introduced and explain in the chapter where they appear. Recording methods, details about participants, and materials will be presented in the method section of the corresponding experiment. All studies received ethical approval from the local ethics committee.

3.1 Behavioural Analysis.

In all the following experiments, we calculated the effects of each condition on responses and, when possible, on accuracy. As shown in the diagrams in Figure 2.1, RT corresponds to the time between the stimulus (for Simple RT) or stimuli (for Choice RT) onset until participant's response. We collected hundreds of responses for each participant, and the number of participants for each experiment ranged from 12 to 30. To obtain a summary description of responses for each condition we take the median RT for each participant and condition and calculate the median of these medians across all participants (we will refer to the resulting data points as mdRT). The reason to use the median is that it is less affected by outliers than the mean (Whelan, 2010). For the Classic RT tasks (Chapter 4, 5 and second part of Chapter 8) the ANOVA is performed on the rate median responses (which is obtained by calculating the median of 1/RT, for each participant and condition), as this appeared to be more close to a Normal Distribution (Spencer & Chase, 1996; Whelan, 2010). For the EXACT Paradigm (Chapter 6, 7, and

first part of Chapter 8), the median responses appeared reasonably Normally distributed, and thus the ANOVA is run directly in the time domain.

For Chapter 4, 5 and 8 (second experiment), the accuracy for each condition is calculated by averaging the percentage of correct responses for each participant and each condition. In Chapter 6 and 7, with the EXACT Paradigm, we will have a direct access to the accuracy value for each single response. In that case, we will use the same approach as with RT, by calculating the median of the median of the accuracy for each response, participant, and condition.

For the experiment in Chapter 5 and onwards we used a point-based experiment, in which participants were asked to earn as many points as possible (points were exchanged for money). For these experiments we analysed the results of sub-groups of participants depending on their performance, by aggregating participants according to their score. The total score for each participant was calculated by summing the score for each condition. The whole participant set was divided into 3 groups based on the 3-quantiles (the lower [33%], the average [33%-66%], the top [66% to 100%] performance group) of the score distribution. Similarly to the aggregated participants, we analysed the three groups by calculating the median of the median response time/accuracy across the participants in each sub-dataset.

3.2 Aggregating Distributions: Vincentizing Method

The individual sample size presented in the experiments of this work is large enough for calculating summary statistics, but too small for distribution analysis of individual participants, which would require at least ~300 responses (Ratcliff & Tuerlinckx, 2002). Therefore, in all our experiments, we decided to aggregate participants' data. The main problem is that participants' responses may easily come from different distributions (possibly, distributions from the same family, with different parameters). A simple data-aggregation may hide the original distribution family. For example, in Figure 3.1 we aggregate two simulated distributions from the same family (a skewed distribution generated by a DDM, similar to what found in real RT data) but with different parameters. The result is a bi-modal distribution, which obviously does not resemble the original DDM distribution.

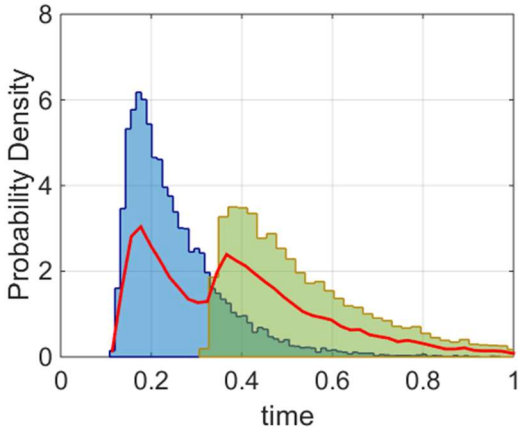


Figure 3.1

Showing how combining two distributions from the same family but with different parameters (blue and yellow) may result in a distribution that does not resemble the original distributions. In this case, we started from two distributions generated by a DDM process and end up with a bimodal distribution, that cannot be produced by a DDM.

To overcome this problem several solutions have been proposed: Sternberg (1969) suggested the use of moments and cumulants to describe distribution shape; Ratcliff and Murdock (1976) assumed an explicit distribution shape, and obtained an aggregated distribution by averaging parameters estimated by individual participants. However, both these methods require a high number of observations for each participant and condition. Furthermore, assuming a particular distribution defeats the aim of model fitting. We used the Vincentizing Method proposed originally by Vincent (1912) and then applied on RT distributions by Ratcliff (1979), a non-parametric approach which retain shape information even with very small datasets (as small as 20 trials per participant). The method simply consists of computing RT quantiles for each participant and condition, and averaging these quantiles over participants to obtain group quantiles. The resulting quantiles represent a hypothetical "average" participant distribution. This method has been tested in Ratcliff work with three distributions (exponential, Weibull and logistic), and has been since then widely used and discussed (Logan, 1992; Andrews & Heathcote, 2001; Genest, 1992; Madden et al., 1999; Penner-Wilger, Leth-Steensen, & LeFevre, 2002). As mentioned in **Section 2.5**, it has been seen that the distributions are sometime Normal in the Rate Domain (which means that they are Reci-Normal in the time domain, see Harris et al., 2014). Thus we tested the Vincentizing method on the Reci-Normal distribution as well, confirming that the method is appropriate for the analysis of this other family of distributions. In order to do that, we simulated a

dataset of 100, 1000 and 10000 trials from 12 different simulated participants. To generate the trials we used a Reci-Normal distribution. The parameters for each individual participant were generated randomly by following realistic parameter values. Once the data were simulated, we calculated the Vincentized distribution (we used 21 quantiles [0 0.05 … 1]) in the time domain and then plot the rate of that distribution. Figure 3.2 shows the result of this analysis for the dataset of sample size 100. It is clear that the Normal distribution shape is retained. We generated similar datasets based on alternative distributions (Exponential Gaussian and Inverse Gaussian), and we verified that these distributions in the rate domain were not similar to a Normal Distribution, as expected.

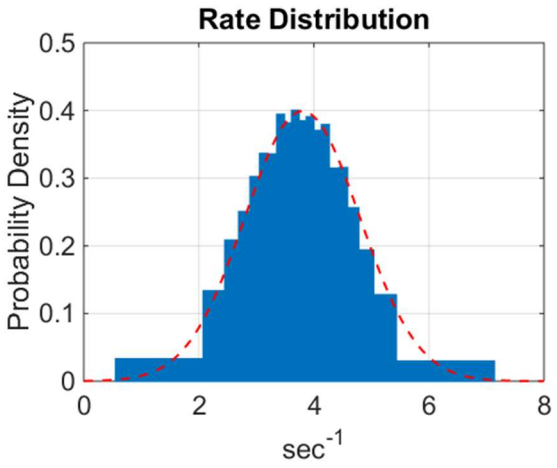


Figure 3.2

The figure show the result (in the rate domain) of applying the Vincentizing method to a set of Reci-Normal distributions. The Normality shape is clearly retained and thus this method seems appropriate for our analysis.

As discussed in **Section 2.5**, numerous studies have observed a certain percentage of "contaminants" response, both short (express saccades) and long (which may be due to memory lapses). These responses may affect the power of model fitting. In order to address this problem, we excluded the extreme quantile bins so that the very short and very long RT would not be taken into account. Therefore we used 19 quantiles [0.05 0.1 0.95].

We concluded that the Vincentized method appears to be a reasonably good approach to aggregate participant's distributions. When comparing Pure/Extended DDM and LATER model we will, therefore, use Vincentized Distributions.

3.3 Normality in the Rate Domain.

The reason to employ the LATER model was to compare a model that generates Normality Distributions with DDM model, in order to test the Rate-Hypothesis (**Section 2.4**). Another way to test Normality is to check directly if the distributions appear Normal in the rate domain. The approach commonly used (e.g. Carpenter & Williams, 1995; Reddi & Carpenter, 2000) is to plot the distribution for each individual participant in the rate domain, generally on a probit plot so that, when the rate distribution is Normal, they sit on a straight line. A Kolmogorov-Smirnov test is generally used for testing for Normality. This approach is valuable when thousands of trials are collected for each individual participant. However, the power of the test decreases dramatically with reducing sample size. In our experiments the number of trials for each participant and each condition is too small to properly estimate the underlying distribution. Aggregating distribution with the Vincentizing method would make impossible to use the Kolmogorov-Smirnov test. Therefore, to test for Normality in the rate domain, we used the method developed by Harris et al. (2014) work, which we call "standardised method".

3.3.1 Standardised Method

The standardised method consists of the following steps: we computed the rate for each response (1/RT) for each participant and condition, then we standardised each dataset into z-scores. At this point all the responses have mean of 0 and standard deviation of 1, which allowed us to collapse all the data together (Harris et al., 2014). By using this method, the resulting distribution is a Normal Distribution if the original ones were normally distributed in the rate domain. In Harris et al. (2014) the truncated distributions are excluded, as they may distort the underlying Normal Distribution when aggregating across different participants. However, we did not find this to be a relevant problem for our studies, as even with the lower contrast level rate distributions appeared far away from zero. The presence of "express" responses, however, complicates this approach. When a few short responses are present in time domain, the corresponding rate distribution is a Normal distribution but with few outliers at very high rate (corresponding to the fast responses in the time domain). When standardising, the high

standard deviation generated by the few high rate responses make the standardised distribution to be unnaturally shrunk (see blue line in Figure 3.3). The few outliers from these "contaminant" responses can thus hide the underlying Normal distribution. One solution would be to select a cut point and eliminate all the responses below that point. However, different conditions may have different cutting points, and the cut point choice may cut some real responses. When collapsing the data, this would result in a left skewed distribution and, again, the real Normal distribution could be distorted. In order to eliminate contaminant data, we developed the following method: for each rate distribution, we checked whether the standard deviation was *X* time higher than the interquartile range (IQR). In any approximately Normal Distribution, this usually implies that few outliers are present. The algorithm keeps eliminating the highest data point from the distributions until the standard deviation is less than *X* time higher than the IQR. The parameter *X* determines the severity of the cutting procedure. By running several simulations (with Reci-Normal distribution, Exponential Gaussian, and Reciprocal Inverse Gaussian Distribution), we checked that this method worked properly in retrieving the original Normal Distribution from a contaminated dataset, and that when applied to a non-Normal distribution did not result in a Normal Distribution. We call this method STD-IQR and show an example in Figure 3.3 (on the left, retrieved distribution from an original contaminated Normal distribution; on the right, retrieved one from a contaminated Exponential Distribution).

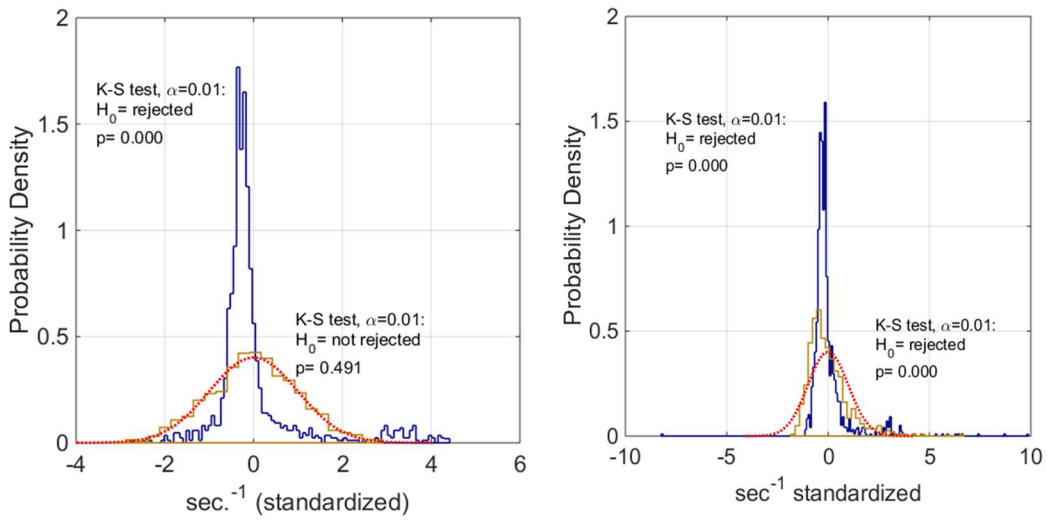


Figure 3.3

In the **left panel**, the blue line corresponds to the result of the standardised method applied to a set of Reci-Normal Distribution obtained from simulated participants. Contaminant responses in the range of 0-200ms were added. The effect of these contaminants is to obtain a "peaked" shape distribution, as shown here. The yellow line shows the same standardised distribution with the STD-IQR method: the Normality shape is retained, as shown by the red dotted line (Normal Distribution fit). In the **right panel** the same method is applied to an Exponential Distribution with contaminants. Again, the standardised method results in a peaked shape. Applying the STD-IQR method does not result however in a Normal distribution (yellow line for the resulting distribution, red dotted line for the Normal Distribution fit), as expected.

3.4 Pure/Extended DDM

In this work, we compare the LATER Model with two important accumulator models used in the decision-making field, the Pure DDM and Extended DDM. We will firstly present the mathematical formulation of these models, then we will describe the fitting and model selection procedure. For a more informal description see **Section 2.3**.

The accumulator part of both the Pure and Extended DDM is described by a Wiener Process:

$$dx = d \cdot dt + s \cdot dW, \qquad x(0) = 0$$

Where $x(t)$ is the accumulator value at time t, dx denotes the change in x over a small time interval dt, the constant drift $d \cdot dt$ represented the average increase of the accumulator towards the correct choice, $s \cdot dW$ represents Gaussian noise (mean 0, variance $s^2 \cdot dt$) (Bogacz et al., 2006, see Figure 2.2). Note that s is a scaling parameter: if changed, all the other parameters can be changed to produce predictions identical to those before the change (Ratcliff & Tuerlinckx, 2002). For the Pure and Extended DDM we set $s = 1$.

3.4.1 Simple RT

For the Pure DDM with Simple RT corresponds to a Wiener Process with one boundary plus some shifting due to sensory-motor processing. The probability of hitting the boundary at time t is found to be an Inverse Gaussian (Wald, 1947). Other processes that are not-decisional, such as stimulus encoding, memory access, motor response, are assembled in the parameter τ, which is added to the accumulator stage, resulting in a RT distributed as a shifter inverse Gaussian:

$$RT(t;a,\tau,d,s) = f(x) = \begin{cases} \dfrac{a}{\sqrt{(2\pi s^2(t-\tau)^3)}} exp\left(-\dfrac{\left(a-d(t-\tau)\right)^2}{2s(t-\tau)}\right), & t>0 \\ 0, & t \le 0 \end{cases}$$

In the Extended DDM with Simple RT we similarly assume a Wiener Process with Gaussian Noise with one threshold, and sensory-motor time added to it. Compared to the Pure DDM, this variant assumes variability across trials for several parameters. The drift rate changes from trial to trial according to a normal distribution with mean d and standard deviation η, and the sensory-motor time τ changes according to a uniform distribution with width s_t. The distribution of RT for the Extended DDM with Simple RT has been found by Patanaik, Zagorodnov, Kwoh and Chee (2014) to be equal to:

$$f(t;a,\tau,s_t,d,\eta,s) = \left(\frac{1}{s_t}\right)\left(G\left(t-\tau+\frac{s_t}{2}\middle|a,d,\eta\right) - G\left(t-\tau-\frac{s_t}{2}\middle|a,d,\eta\right)\right)$$

Where

$$G(t;a,d,\eta,s)$$
$$= \begin{cases} 1-\Phi\left(\dfrac{a-dt}{\sqrt{t^2\eta^2+s^2t}}\right) + ex\,p\left(\dfrac{2ad+2a^2\eta^2}{s^2}\right)\Phi\left(-\dfrac{as^2+2at\eta^2+s^2dt}{s^2\sqrt{t^2\eta^2+s^2t}}\right), & t>0 \\ 0, & t \le 0 \end{cases}$$

3.4.2 Choice RT

The Pure and Extended DDM for Choice RT are dissimilar to the Simple RT version in that they present two boundaries instead of one, reflecting the possibility of choosing one amongst two possible alternatives. For the Pure DDM, the RT distribution for the correct choice boundary (that is, the boundary in the direction of the mean drift) is equal to (Ratcliff & Smith, 2004b):

$$RT(t,r=1;a,\tau,b,d) = \frac{\exp\left(ad(1-b) - \dfrac{d^2(t-\tau)}{2}\right)}{a^2}\ f\left(\frac{t-\tau}{a^2}\middle|b\right)$$

Where $r=1$ indicates the correct responses, and $f(u|b)$ is defined as

$$f(u|b) = \pi \sum_{k=1}^{+\infty} k \exp\left(-\frac{k^2\pi^2 u}{2}\right) \sin(k\pi b)$$

where usually the infinite summation is evaluated until convergence below an error bound ($\varepsilon = 10^{-10}$). Note that this is a defective distribution, meaning that its area is not equal to one, but to $p(r = 1|t, a, \tau, b, d)$. The distribution of responses to the lower boundary (incorrect response) is

$$RT(t, r = 0; a, \tau, b, d) = RT(t, r = 1; a, \tau, 1 - b, -d)$$

which is also a defective distribution. The sum of the area of the two distributions is equal to unity.

The distribution of RT for the Extended DDM has no analytic formulation but has to be numerically computed (Ratcliff, McKoon, Van Zandt, 1999; Ratcliff, Tuerlinckx, 2002).

3.5 LATER Model

In the LATER model, a linear accumulator with drift rate r starts from S_0 and rises until reaching S_t. The drift rate is normally distributed across trials, with mean μ and standard deviation σ (Carpenter & Williams, 1995). Therefore, the time for responding is given by $(S_t - S_0)/\Phi(\mu, \sigma)$, where Φ is the Normal probability density function. The density function for this random variable, which corresponds to the distribution of RT, can be easily found by considering that the density of an inverse random variable $Y = 1/X$, where X is continuous with density function $f(t)$, is equal to $\frac{1}{t^2} f\left(\frac{1}{t}\right)$ and the density of a random variable $Y = cX$ where c is a constant and X is continuous with density $f(t)$ is equal to $\frac{1}{|c|} f\left(\frac{i}{t}\right)$. By setting $(S_t - S_0) = a$ (threshold distance to starting point) and applying the aforementioned properties, we obtain

$$RT(t; a, \mu, \sigma) = \frac{a}{t^2} \Phi\left(\frac{a}{t}, \mu, \sigma\right)$$

Here, we consider a slightly different version of this model: negative rates have a difficult interpretation, as they imply negative time of response (or infinite type of responding, depending on the

model interpretation). Therefore, we will substitute $\Phi(x;\mu,\sigma)$ with a left truncated (at 0) Normal Distribution, which probability density function is:

$$RT(t;a,\mu,\sigma) = \frac{\exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\left(1-\Phi(-\mu/\sigma)\right)\sigma t^2}$$

where Φ indicates the Normal standard cumulative density function*. During this work, we will use this version of the LATER model. In term of Rate-Hypothesis, this implies that the computation of the decision process is done on the rate domain, according to a *truncated at 0* Normal Distribution.

3.6 Decision Rules and Optimum Decision Time with Exponential $ACC(DT)$

The mathematical formulation of each decision rules is shown in the left column of Table 3.1. D_{tot} corresponds to all the trial time that is not decision time, that is: the time from a trial to another (called D, for delay, and may correspond to the time from the participant response until the starting of the next trial, or to the beginning of the trial until the onset of the stimulus, etc.) plus the sensory-motor component τ. See Figure 2.1 for a diagram of the relationship between RT, DT, τ and D_{tot}. The parameter q is an "emphasis on accuracy" parameter (Bogacz et al., 2006): by increasing it, participants will weigh more accuracy than speed. For RR_m, the q parameter can be more conveniently interpreted as the subjective punishment/reward ratio. See **Section 2.6** for more details.

In Chapter 6, 7 and 8 we use an exponential formulation of $ACC(DT)$ (shown on top of Table 1.1) that we incorporated into the experimental task. The $ACC(DT)$ defines the speed accuracy function: how accuracy increases with increase in decision time (DT), where accuracy is defined as the

* Note that, however, with most experimental conditions used in this work, the difference between these two versions is very small. When the parameters were fitted to Experiment 1, for example, the lowest μ was ~0.8 and σ was found to be 0.15, resulting in a probability of negative rate of 4.8213e-08.

probability of a correct response. In this equation (see top of Table 1.1), λ corresponds to how fast the accuracy increases in time, and α is the accuracy at the beginning of the trial, that is $ACC(0)$. For each decision rule, given a specific formula for $ACC(DT)$, we can calculate the optimum by setting the derivative of each decision rule equal to zero and solve for DT. The solution for RR and RR_m were only found in terms of the $Lambert\ W$ function, defined to be $W(z)\,exp[\,W(z)] = z$ (see Appendix). The optimum decision time DT^* given the exponential $ACC(DT)$ are shown in the right column of Table 3.1

Table 3.1 Decision Rules and Optimum Decision Time for exponential $ACC(RT)$

Decision Rules	Optimum decision time (DT^*), using $ACC(DT) = 1 + (\alpha - 1)\,exp(-\lambda DT)$
$BR = -\left[DT + q\left(1 - ACC(DT)\right)\right]$	$DT^*_{BR} = -\dfrac{ln\left(\dfrac{1}{\lambda q(1-\alpha)}\right)}{\lambda}$
$RR = ACC(DT)/(DT + D_{tot})$	$DT^*_{RR} = -D_{tot} - \dfrac{W_{-1}\left(\dfrac{exp(-\lambda D_{tot} - 1)}{\alpha - 1}\right) + 1}{\lambda}$
$RA = RR - \dfrac{q}{D_{tot}}\left(1 - ACC(DT)\right)$	*No Explicit Form*
$RR_m = \dfrac{ACC(DT) - q\left(1 - ACC(DT)\right)}{DT + D_{tot}}$	$DT^*_{RR_m} = -D_{tot} - \dfrac{W_{-1}\left(\dfrac{exp(-\lambda D_{tot} - 1)}{(\alpha - 1)(q + 1)}\right) + 1}{\lambda}$

3.7 Distributions: Fitting Routine and Model Selection: QMLE, AIC

When the Pure/Extended DDM or the LATER models are fitted to experimental data, we can make several assumptions about parameters restrictions. For example, let's assume we have a Simple RT task where we use two different stimulus contrasts, dim and bright, and we want to test whether this has an

effect on some parameters of the Pure DDM, for example the drift rate d. In this case we want to compare two *versions* of the same models, one with free d for the two contrast conditions, and the other one with fixed d. The version with free d has 4 parameters: a, τ, d_1, d_2, whereas the version with fixed d has 3 parameters: a, τ, d. Each model version, together with the empirical (Vincentized) distribution, is given as an input to a Model Evaluation Routine. This minimization function is an extension of *fminsearch.m* in MATLAB, which employs the Nelder-Mead algorithm. This extension included the possibility to specify parameter bounds, which allows faster convergence. For the Pure and Extended DDM, we specified 0 as the lower bound for all parameters (no higher bounds were used). The minimization routine finds the parameters that minimize the summed Quantile Maximum Likelihood Function across all conditions (Heathcote et al., 2002; see also reply to Speckman & Rouder, 2004). This method is an extension of the Maximum Likelihood Estimation (MLE) for Vincentized Distribution: QMLE has been proved to be typically a little less biased and much less reliable than other commonly used method, such as Continuous Maximum Likelihood and MLE (Brown & Heathcote, 2008). In the QMLE, we seek to maximize the quantity

$$L(\theta|T) \propto \prod_{j=i}^{m} \left(\int_{\hat{q}_{j-1}}^{\hat{q}_j} f(t,\theta)dt \right)^{N_j}$$

with respect to the transformed data $T = (N, \hat{q})$, $\hat{q}$ is a vector of quantile estimates and N is a vector of counts of the number of RTs that occur in each interquantile range. In order to avoid overflow, instead of maximizing $L(\theta|T)$ we minimized $-\ln L(\theta|T)$. Note that, generally, by using the vector count N, we are automatically adjusting the likelihood for error and correct distribution (errors are generally less than correct responses, and they have to be weighted accordingly). The function f is the probability density function for the model we are evaluating (in this example, the Pure DDM probability density function, shown above).

The first version of this example, with two drift rates, is more flexible and thus will obviously fit at least as good as the second model. However, we desire not only to obtain a model with a good fit, but

a parsimonious model. To do that, we use a metric that penalized the model according to the number of parameters used: for doing that we use the Akaike information Criterion (AIC), computed as

$$AIC = 2k - 2\ln\left(\sum_{c=1}^{n} L(\theta_c|T_c)\right)$$

Where k represents the number of free parameters, and the second part of the equation computes twice the natural logarithm of the sum of the likelihood values across all n conditions, with the relative θ parameters and T dataset for that condition. The preferred model is the one with the minimum AIC value. AIC gives a measure of the information lost for a model, dealing with the trade-off between goodness of fit, $L(\theta|T)$, and complexity of the model (Burnham & Anderson, 2002). It is useful for comparing models (or model variants) with different number of parameters, as it penalizes for the number of parameters.

Note that by using this minimization routine we will obtain the AIC value, the summed likelihood across all conditions and the likelihood for each condition, so more fine model comparison for particular conditions in terms of goodness of fit can be also made.

3.8 Data points: Fitting Routine and Model Selection: WLS, BIC

In some cases we are not going to fit distributions but data point

s, for example to check whether the change in mdRT across experimental conditions followed the prediction of a decision rule. The Fitting Routine used is exactly the same as the distribution procedure: we first established what are the free parameters across experimental conditions, generating if necessary different versions of the same model (which has now a functional relationship between experimental condition and observations). Each model version, with the observation dataset and standard deviation is given as an input to the Model Evaluation Routine. The modified $fminsearch.m$ find, for each model version, the parameters that minimize the summed Weighted Least Square (WLS) error across all condition. This is very similar to the Minimum Least Square Method, but takes into account the

standard deviation for each data point (which can calculate as being the standard deviation across median responses for each participant). The WLS is defined as

$$WLS = \sum_{i=1}^{n} \frac{\left(y_i - f(x_i|\theta)\right)^2}{\sigma}$$

where n is the number of observation, (x_i, y_i) is a dataset of observations, θ is the set of parameters, σ is the standard deviation for each data point. With this method, the deviation from an observation with low standard deviation will have a higher error than the same deviation in an observation with higher standard deviation.

Once the parameters have been minimized for each model version, we proceed to the model selection, by selecting parsimonious model and thus penalizing for the number of parameters. To do that, we apply the modified Bayesian Information Criterion (BIC, Schwarz, 1978):

$$BIC = -2ln\left(\frac{1}{n}\sum_{j=1}^{C} WL_{\ j}\right) + \frac{(k+1)ln(n)}{n}$$

where C indicates the total number of condition, WLS_j is the WLS computed for a particular condition j, n the total number of observations, and k the number of parameters. Like AIC, the BIC compute the trade-off between the goodness of fit and the model complexity (Schwarz, 1978), and the preferred model is the one with the lower BIC.

3.9 Response Signal Analysis: STARS

We introduce the analysis of RT as response signals, in which each response is associated with the time it has been produced. We performed a Regime Shift Detection Analysis, usually performed in the field of ecosystem analyses to detect rapid change or reorganisation in the ecosystem. We used this measure to detect a change in the internal strategy used by the participants. Amongst several ways to analyse discontinuities in time-series (Easterling & Peterson, 1995; Lanzante, 1996) we chose the Rodionov (2004) sequential t-test analysis of regime shifts (STARS), an automatic procedure that does

not require data pre-processing, and detects shifts and provide the magnitude of the shifts even for short time series. As this technique is uncommon in the field of Psychology, we provide a short summary here and refer to Rodionov (2004) and Rodionov and Overland (2005) for further details.

The method consists of defining a time window of size l, and establish whether a data point is significantly different from the mean of the current regime by using the Student's t-test: $diff = t\sqrt{(2\sigma^2)/l}$ where l is the length of the time-series window (we set this equal to 7), σ_l is the averaged of the standard deviation calculated for each time-series within the window, and the critical level of the t-test was set to 0.05. A data points deviating more than $diff$ to the mean of the current regime is marked as a potential change of shift. Subsequent observation are used to confirm this hypothesis by calculating the regime shift index

$$RSI = \sum_{i=c}^{c+m} \frac{x_i^*}{l\sigma_l}$$

where $m = 0, \ldots, l-1$, c is the index of the potential regime shift. RSI represents the cumulative sum of normalized deviations x_i^* from the average value of the current regime. If from c to $c+m$ (that is, for the window of size l starting from the first deviation point c), the RSI remains positive, then the point c is marked as a new regime shift, and the search for a new regime shift starts at $c+1$.

Chapter 4

Intensity and Foreperiod Effect on Simple RT

4.1 Introduction

In this chapter, we analyse the general effect of foreperiod on RT by using a Simple RT design and compare three models of decision making: the Pure DDM, Extended DDM and LATER model. In doing that, we will provide some qualitative evidence for some of the decision rules considered in this work.

4.1.1 Literature on FP

The foreperiod (FP) corresponds to the time between the disappearance of the warning signal and the appearance of the stimulus and provides a temporal frame of reference for the participant to prepare for responding (Niemi & Näätänen, 1981). The relationship between FP and RT has been analysed for decades (Luce, 1986), but more recently has been largely neglected. Most of the studies on FP uses Simple RT tasks (the literature for FP and Choice RT will be analysed in Chapter 5 when we will use a Choice RT task). During each block, the FP may be maintained constant or it may be varied according to a certain distribution. With variable FP (FP generated from a random variable which parameters vary across different blocks), the RT are generally longer than with constant FP (e.g. Bevan et al., 1965), and RT decreases monotonically with actual FP duration (Aiken & Lichtenstein, 1964; Drazin, 1961; Klemmer, 1956; 1957; Nickerson, 1965a,b). When variable, the FP is drawn from a particular distribution. The two most used distributions for the variable FP are the uniform and the exponential distribution. In this work, we used the exponential distribution, as it does not provide any information about the appearance of the stimulus as time flows (Luce, 1986, Elithor & Lawrence, 1955). This means that the probability of the stimulus appearing at time t, given that it had not appeared in the past, is the same for every t. However, it turns out that responses are not remarkably affected by the FP distribution. For both exponentially and uniformly distributed FP, it has been found that RT increases with higher

mean FP (Granjon, Requin, Durup & Reynard, 1973; Nickerons & Burnham, 1967; Karlin, 1959, Niemi & Näätänen, 1981) and RT decreases with actual FP duration (Näätänen, 1970; Hermelin, 1964). We suggest that the lack of difference between uniformly and exponentially distributed RT resides in a misconception about what the participant is estimating during a simple RT, which is not the probability that the stimulus occurs at each point in time (the hazard function) but the probability that the stimulus has appeared at some time in the past (the cumulative density function, $P(T < t)$). In fact, in a Simple RT participants are usually awarded for responding to the stimulus onset as fast as possible. The participant continuously receives information about the status of the stimulus on the monitor screen (that is, if a stimulus is in fact there or not), but this information is weighted with the probability that the stimulus onset happened already sometime in the past (which is described by the cumulative density of the FP distribution). By taking this aspect into account, we can easily see how the RT for the exponential and uniform distribution presents the same patterns, since both distributions have a monotonically increasing cumulative density function. This aspect of RT task will be briefly analysed in the following experiment.

It is also interesting to consider how the number of anticipations varies with different FP. Unfortunately, data for anticipation are not always presented, probably because in these tasks anticipations are usually less than 5% of total responses and no analysis is possible (e.g. Green, Smith & von Gierke, 1983; Bernstein, Rose, & Ashe, 1970).

4.1.2 Effect of FP on DDM and LATER Model.

By using the Drift Diffusion Model (DDM) we can infer what part of the decision mechanism is affected by varying the FP. This has been mostly explored within the context of Choice RT (Bausenhart, Rolke, Seibold & Ulrich, 2010; Seibold et al., 2010; Jepma, Wagenmakers, &Nieuwenhus, 2012), but the same paradigm can be applied to Simple RT as well. The general question is: what parameters of the DDM are affected by varying the FP? Three different possibilities have been identified:

1) FP affects the drift rate (the rate of accumulation of information, parameter d): higher temporal certainty (obtained with shorter FP) may increase participant's attention towards the stimulus, and

therefore increase the accumulation rate (Grosjean, Rosenbaum & Elsinger, 2001). High drift rate corresponds to *faster* RT and *higher* accuracy (lower error rate).

2) By affecting the sensory-motor component of response time (parameter τ): participant may be more prepared to the stimulus onset with shorter FP (Rolke & Hofmann, 2007). Large τ corresponds to *slower* RT but no difference in accuracy (or error rate). This was found by Seibold et al. (2010) in an experiment with different FP conditions and different proportion of catch-trials and nogo-trials, and the result was again confirmed by Jepma, Wagenmakers and Nieuwenhus (2012) in a 2AFC letter recognition task with 2 FP conditions (0.3s and 2.5s). Surprisingly, they also found a decrease in accuracy with shorter FP in the condition where accuracy was stressed, which usually pinpoint to a difference in boundary separation instead (next paragraph). Their explanation for this drop in accuracy was that participants were starting sampling too early, therefore increasing RT but decreasing accuracy (see also Laming, 1968).

3) by affecting the threshold separation (parameter a), that is the stopping rule of the accumulator process, which would result in *slower* RT and *higher* accuracy (lower error rate). This is particularly relevant in terms of optimization. In fact, three out of the four decision rules we are considering in this work (RR, RR_m and RA, but not BR) are function of the non-decision time (D_{tot} , the delay between participants' decisions). In this study, by increasing the FP, we are in fact increasing the total delay across responses. If participants are indeed employing one of the 3 decision rules, then an increase in FP should correspond to an increase in the threshold separation (as explained in **Section 2.6**, see also Bogacz et al., 2006).

We test these three hypotheses for the first time with Simple RT. As explained in **Section 2.3**, although in a Choice RT different parameters affect both RT and accuracy, in a Simple RT only one choice is possible (the DDM has only one boundary so that it cannot produce incorrect responses). However, we will take into account the number of anticipations (responses before the stimulus onset), which will serve as a rough indication of accuracy.

As for the LATER model, to the extent of our knowledge, it has never been applied to manual Simple RT. The only two parameters that could possibly change are a (that is, the threshold distance) and the rate of linear increment r. According to Carpenter (2004), the contrast should affect the parameter r, but the effect of FP on LATER parameters has not been investigated.

4.1.3 Aim of Experiment 4.1

The study in this Chapter was designed to investigate three main topics:

1) Investigate the effect of FP in terms of RT and accuracy (by using the number of anticipations) in a Simple RT. As discussed above, the accuracy level is almost never reported for Simple RT, and the data are usually excluded from the analysis. However, this data can be useful to make predictions about the underlying mechanism taking place when manipulating the FP.

2) Investigate the FP effect within the framework of the DDM and LATER model. The DDM has been applied to the FP effect only within the context of Choice RT. Investigating what parameter changes for Simple RT would either confirm previous findings or suggest a different mechanism for the two experimental designs. We compare a version with free threshold, free sensory-motor component, and a combination of both parameters free across FP conditions. The Extended DDM has been recently applied to Simple RT (Ratcliff & van Dongen, 2011; Ratcliff & Strayer, 2014). However, the historical reason to use the Extended version of the DDM is to take into account error responses (Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004). Neither the Pure DDM nor the Extended DDM can take into account error responses (anticipations) in Simple RT, so using the Extended DDM in place of the simpler Pure DDM may be unnecessary. Note that LATER model and Pure DDM have the same number of parameters, whereas the Extended DDM model has, for the Simple RT paradigm, two additional parameters (st and η). As discussed above, LATER has been successfully used for modelling saccading responses to stimulus onset, but no comparison between these three models has been performed in the past. Plus, in recent years there has been a surge in the importance of the Pure DDM even for Choice RT (Bogacz et al., 2006; Balci et al., 2011; Moran, 2014; Simen et al., 2009) due to its optimality property, so the need for a comparison is compelling.

3) Investigate the rate distributions for the Simple RT task. The observation that distributions of responses are normal in the rate domain (Harris et al., 2014) for Choice RT suggests that what is actually computed by the observer is the rate of response, and therefore the DDM (or any similar rise-to-threshold model) could be a mechanism to convert from time to rate (this is what we called the Rate-Hypothesis, **Section 2.4**). The analysis of response distributions in the rate domain for a Simple RT has not been performed in the past.

EXPERIMENT 4.1

4.2 Methods

Participants. 12 participants (5 females, 7 males) took part in the experiment. All participants had normal or corrected-to-normal vision and no known neurological conditions.

Materials. All testing were conducted under constant levels of illumination. The stimuli were presented binocularly on a computer monitor on a PC running Windows XP, using the software E-Prime (version 2.0.10). Participants were positioned 57cm from the monitor. The stimuli were luminous circles (2cm, ~2°) shown on a grey background (0.31 cd/m^2). The five levels of luminance followed an approximately geometrical series (0.42, 0.71, 1.21, 2.06, 3.50 cd/m^2). These values were converted to Michaelson contrast (defined as $\frac{L_s - L_b}{L_s + L_b}$, where L_s is the luminance of the stimulus and L_b is the background luminance) resulting in 0.15, 0.39, 0.59, 0.74, and 0.84. We used three different FP conditions: short, medium and long. Each FP condition was kept constant for each of the three blocks. During each block, the 5 possible stimuli were presented randomly. Each stimulus was repeated 60 times. At the beginning of the block, the participants were dark adapted, and pauses were made between blocks of different FP. Each FP condition consisted of a minimum waiting time (0.4s) added to a number of millisecond drawn from an exponential distribution. Foreperiod values higher than 15s were excluded. The mean of the exponential distribution was 0.2s for the short condition, 0.6s for the medium condition and 2s for the long condition, leading to three different FP conditions with mean of, respectively, 0.6s,

1s and 2.4s. The mean FP (mFP) time was kept constant throughout each block. The mFP conditions were counterbalanced across participants. We did not use a timeout.

Procedure. A white cross (2 cm) was presented on a black background at the centre of the screen. This was used as a fixation point. After 1 second, the cross disappeared and a white dot (1mm) appeared at the centre of the screen and remained at the centre during the whole trial. The disappearance of the white cross constituted the warning signal for the beginning of the trial. After the FP time, the stimulus appeared and stayed on the screen until the participant pressed the space bar (Figure 4.1). The participants were asked to press the spacebar as soon as the stimulus was detected. After the spacebar was pressed, the stimulus disappeared. If the participant pressed the spacebar before the stimulus appeared (anticipation response), a negative auditory feedback was provided and the trial started again; otherwise, a positive auditory feedback was provided. After the response, there was a delay of 1 second before the starting of the next trial. Each luminance level was randomly presented within a session. We recorded 60 RTs for each luminance level. Thus, we recorded 300 non-anticipation RT for each FP condition and a total of 900 non-anticipatory RT for each participant. A training session consisting of 20 trials with a short FP was performed by each participant before the beginning of the experiment.

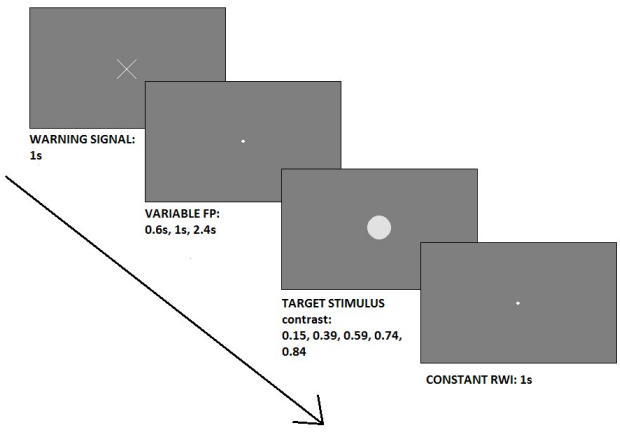


Figure 4.1

The flow of events in the study.

Analysis. The analyses were performed as illustrated in Chapter 3. In this study, we also include an analysis of the effect of FP duration (not to be confused with the mean FP for each block) on the RT.

In order to do this analysis, we proceeded as follows: we divided the FP distribution in 20-quantiles [0.05 0.1 … 1]. We associated each quantile with the RT included in the quantile range, so that for each quantile we obtained approximately the 5% of total responses for that condition. The average RT for each quantile was computed, resulting in 20 RT bins. The resulting average dataset can give us a visual representation of the relationship between FP duration and RT. To better understand the severity of this initial relationship, we fitted an exponential function in the form of $RT = \alpha + \beta\exp(\gamma t)$ where t corresponds to the midpoint of each bin. The function is evaluated only until the mean FP for each condition. To find the α, β and γ parameters we used a minimization routine on R^2 (we do not follow the convention detailed in Chapter 3 as, in this case, we are not going to compare different models, and R^2 gives a direct measure of goodness of fit).

4.3 Results

4.3.1 Behavioural Responses

The anticipation responses (2.7% of the total dataset) will be analysed in a subsequent section. For the remaining analyses, we excluded the anticipations. Figure 4.2 shows the median of median reaction times (mdRT) grouped across participants. The bar indicates ± 1 standard error. As expected, the mdRT decreased with increasing contrast. Using the median rate values (see Methods) on a repeated measure ANOVA revealed that the effect of contrast was significant, as expected ($F_{4,44} = 17.69, p \cong 0$). The effect of mFP was also significant ($F_{2,22} = 8.794, p \cong 0$), but no significant interaction was found between FP and contrast levels ($F_{8,88} = 1.576, p = 0.144$). Post hoc pairwise comparisons with Bonferroni correction revealed that each contrast level was significantly different from one another at $\alpha = 0.05$, apart from the first and the lowest and the second to lowest contrast condition ($p = 0.101$). All three mFP conditions were significantly different at $p < 0.01$. By looking at the figure, it is clear that increasing FP resulted in an increase of mdRT, whereas contrast decreased mdRT in a Piéron's-like fashion (**Section 2.9**).

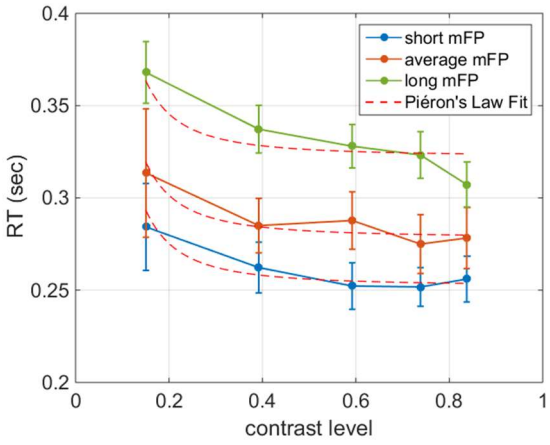


Figure 4.2

Resulting mdRT for the three different mFP conditions and the five contrast levels. Increasing contrast appear to decrease mdRT according to Piéron's Law (red dashed line), whereas the effect of the mFP condition appears to be a shifting of the RT curve.

We fitted Piéron's Law ($mdRT = \gamma + kI^{-\beta}$) on the aggregated data by assuming that k, β or γ (or any combination of those, or none of them) were free to vary across FP conditions, and keeping fixed the remaining parameters. This resulted in testing 7 different model variants (Table 4.1). We compared the models by using the modified BIC for weighted least square (see **Section 3.8**). The best fitting model resulted to be the one with γ free across FP conditions (df 5). The BIC values for all models are shown in Table 4.1, and Table 4.2 shows the estimated parameters for the preferred model. We also fitted the data to the Luminance levels instead of the contrast level, which confirmed that the data followed Piéron's Law and the best model was the γ free.

Table 4.1 Fitting Piéron's Law Models

Model	BIC	df
k free	-4.58574	5
β free	-2.68943	5
γ free	**-6.02601**	**5**
k and β free	-4.10255	7
k and γ free	-5.93085	7
β and γ free	-4.32684	7
k, β and γ free	-4.81213	9

Table 4.2 Parameters for the best model

Model 3: γ free	
BIC	-6.026
ΣWLS	0.0123
$\boldsymbol{k}$	0.001
$\boldsymbol{\beta}$	2.041
$\boldsymbol{\gamma_{mFP=0.6s}}$	0.252
$\boldsymbol{\gamma_{mFP=1s}}$	0.279
$\boldsymbol{\gamma_{mFP=2.4s}}$	0.323

Intuitively, the γ parameter is connected with the sensory-motor component τ for the DDM, but we will see how things can be more complicated than that in future Sections. The linear dependency between contrast and FP suggests, furthermore, that the two experimental factors may affect two different components of the decision process. By using DDM fitting, we will be able to quantify more precisely these hypotheses.

4.3.2 Fast Responses and Anticipations

It is a common observation to find in these datasets a series of very fast responses which are usually assumed to be part of a secondary process, called "express saccades" (Fisher & Boch, 1983; Fischer & Ramsperger, 1984) in the literature of visual decisions (**Section 2.4**). In fact, we found a series of very fast responses in our distributions. The left panel of Figure 4.3, shows an example of a collapsed distributions across all participants, for the conditions with shortest FP and lowest contrast. This figure illustrates the presence of a small set of responses in the fast range that may not be part of the main distributions. Figure 4.3 (right panel) shows the percentage of fast responses (based on the left panel of Figure 4,3, we used $RT < 0.2s$) for each condition. There was a strong and significant ($F_{2,22} = 29.99, p < 0.001$) effect for mFP, but not for contrast. The number of fast responses may be connected with the predictability of the stimulus occurrence. This means that, as FP variance increase, we expect less and less fast responses, which is exactly what we observed. Note how the number of fast responses is extremely high in the shortest mFP condition, with the mean of 0.6s (0.4s of minimum waiting time

+ 0.2s of exponentially distributed time). The very low variability for this condition may have made possible for participants to treat this condition as a constant FP, thus anticipating the appearance of the stimulus (for this reason, in the next experiment, we will not employ the minimum waiting time of 0.4s). As we will see in **Section 4.3.5**, this high number of fast responses may mask the underlying distribution, and they may be the product of a different mechanism based on the flow of time (Ratcliff, 1993; Ratcliff & Tuerlinickx, 2002; Ulrich & Miller, 1994, see Discussion)

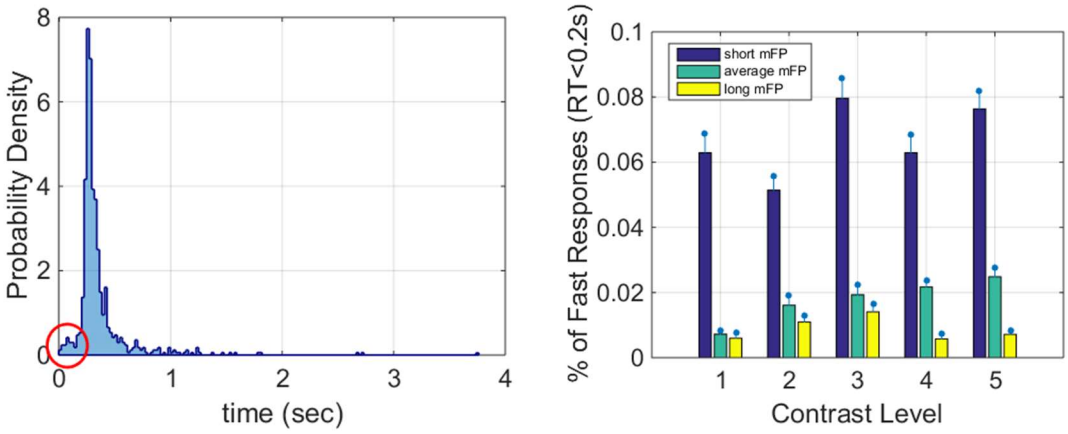


Figure 4.3

(Left panel) Aggregated RT distribution for the short mFP condition and the lowest contrast level condition. The red circle indicated a cluster of very fast responses that appear to be separated from the main distributions and they may due to a different process, e.g. time estimation process. **(Right panel)** The average percentage of fast responses ($RT < 0.2s$) for each condition. The blue line represents 1 standard error. There seems to be a non-linear decrease in the amount of fast responses as the mFP increases, with most of the fast responses obtained in the shortest mFP duration.

It is also possible that the number of fast responses related to the number of anticipations (responses before the stimulus onset). The analysis of anticipations is particularly relevant because it underpins different effects in the underlying decision mechanisms, as we will see when fitting the DDM. We calculated the average percentage of anticipations for the three mFP conditions and the five contrast conditions across participants (Figure 4.4, top panel). Note how, except for the fourth contrast condition, anticipations were always decreasing with higher mFP. We performed the same analysis by collapsing together across mFP and contrast conditions separately (see Figure 4.4 bottom panels, bars indicate 1 standard deviation). There did not seem to be any pattern for different contrast level, whereas the percentage of anticipations seemed to decrease with decreasing mean FP, even though this difference

was not significant ($F_{2,33} = 0.35, p > 0.5$). We will return on the meaning of this analysis in the following Sections.

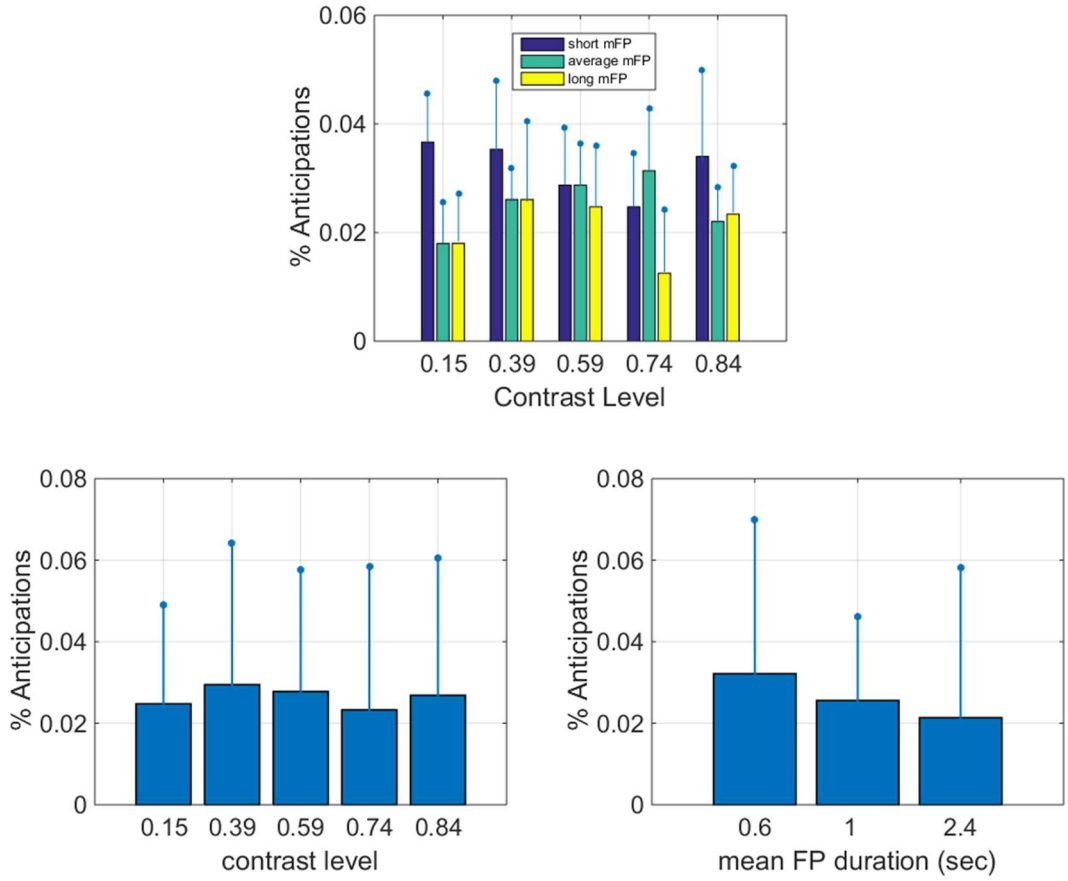


Figure 4.4

(Top panel) Percentage of anticipation for each condition, with the blue line indicating one standard error. **(Bottom panels)**. Percentage of anticipation aggregated across contrast level conditions and mFP conditions. The first does not appear to have any relevant effect, whereas increasing the mFP appear to decrease the percentage of anticipations.

4.3.3 Foreperiod Duration and RT

The effect of FP has been mostly investigated in terms of the effect of mFP length. However, it is interesting to understand what is the effect of the FP for each individual trials. The main point of this analysis is to infer how the information obtained from the flow of time itself before the appearing of the stimulus affects the related response (Näätänen, 1970). The results are shown in Figure 4.5.

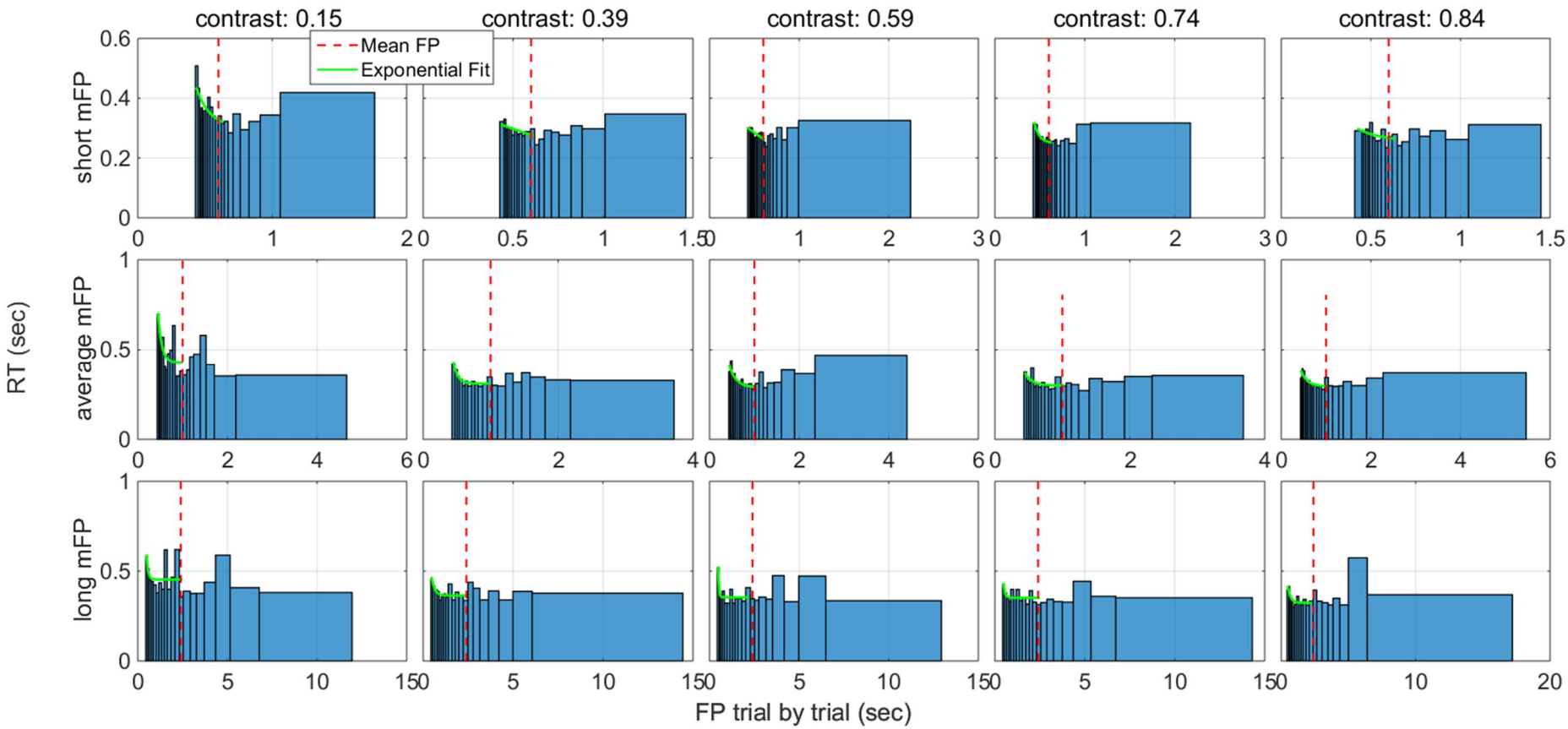


Figure 4.5

Relationship between FP duration and RT (not to be confused with the relationship between the mFP condition and RT) for each condition. Note how the panels are scaled differently. The red dashed line indicates the mean FP time for each condition. There is an initial decreasing relationship that appear to be stronger for the short mFP condition, and that we tried to describe with an exponential function (green line).

By visual inspection, it appears that there is an initial decreasing relationship between FP duration and RT, that saturates at around the mean of the FP condition (represented by a dashed red line in Figure 4.5). The relationship appears to be less clear after the mean FP. Generally, the slope of the decreasing relationship seems to depend on the contrast used, with lower contrast conditions having a steeper decrease before settling to a steady state. To better understand this relationship, we tried to fit an exponential function in the form of $\alpha + \beta\exp(\gamma t)$. Even though by visual inspection the fitting appeared to capture the observatiosn (green line in Figure 4.5), the resulting R^2 values were low, indicating a poor fit (Table 4.3) and it did not seem to be any relevant pattern between conditions and estimated parameters, apart from the value of β (the asymptotic value) which unsurprisingly reflects the general findings about mdRT (highly mdRT with higher mFP).

Table 4.3. R^2 values for the exponential fitting in RT-FP relationship

R^2	Contrast=0.15	Contrast=0.39	Contrast=0.59	Contrast=0.75	Contrast=0.84
Short mFP	0.494	0.366	0.580	0.775	0.193
Average mFP	0.518	0.834	0.639	0.332	0.544
Long mFP	0.178	0.567	0.715	0.358	0.617

Even if it did not match an exponential fit, we believe that the initial decreasing relationship is likely connected with the motor preparedness depending on the flow of time and, as we claimed in **Section 4.1.1**, it confirms that participants are estimating the cumulative distribution function of the FP so to get more prepared as time flows towards the mean of the FP distribution.

4.3.4 Model Fitting

By fitting the drift diffusion model we can have a more precise idea about what component was affected by mFP manipulations. In particular, it is relevant to understand whether the τ (sensory-motor component), the α parameter (threshold distance) or the d parameter (drift rate) was affected by changing mFP (**Section 4.1.2**). We fitted the model to the one-boundary version of the Pure DDM and to the Extended DDM to the aggregated, Vincentized dataset (**Section 3.4**). We use 5 different variants of the model. For each of the following models, we set the parameter d to be free to vary for each

contrast condition, as it is well established that stimulus strength affects the drift rate (e.g. Ratcliff & Rouder, 1998; Donkin & Van Maanen, 2014). The models are presented in Table 4.4 (the bold Variant is the one which resulted in the lowest AIC value). A part for the d parameter, all the non-specified parameters were kept fixed across experimental conditions. The same variants were used for the Extended DDM, with the addition of parameters st and η (variability in sensory-motor component and variability in the drift rate across trials), which were assumed to be fixed across experimental conditions.

Table 4.4. Model variants used for the DDM

Variant 1	Both a and τ fixed across for all conditions
Variant 2	**a free for mFP conditions**
Variant 3	τ free for mFP conditions
Variant 4	τ and α free for mFP conditions
Variant 5	d free for mFP condition (and contrast)

The LATER model was analysed in a similar fashion. To the extent of our knowledge, this was the first fitting of a LATER model to a Manual Simple RT task, so we did not know precisely what parameters were more likely to be affected by mFP conditions. On the other hand, it seemed clearly established (Carpenter, 2004) that contrast affected the μ parameter, so that we set this parameter to change across contrast conditions. The variants we used are shown in Table 4.5.

Table 4.5. Model variants used for the LATER model

Variant 1	Both a and σ fixed across for all conditions
Variant 2	**a free for meanFP conditions**
Variant 3	σ free for meanFP conditions
Variant 4	μ changing for meanFP (and contrast)

Note that one important difference between these models is that the LATER model does not have a τ parameter (which, if added, would not generate a Normal Distribution anymore, which is the main reason behind using the LATER model), and the σ parameter is left free to vary.

The fitting of the Pure and Extended DDM resulted in different preferred models. For the Pure DDM, the lowest AIC is obtained by Version 4 (a and τ free), whereas the preferred Extended DDM and LATER model was Version 2 (which corresponded, for both, to the a parameter free across mFP conditions). The lowest AIC value was obtained by the Extended DDM. The AIC values for all models are presented in Table 4.6. As the Table shows, the difference between the two DDM variants is very small, whereas the AIC value for the LATER model seems remarkably higher than the other two models (recall that lower values mean that the model is preferred according the AIC metric).

Table 4.6. Fitting results for DDM and LATER model

Pure DDM			**Extended DDM**			**LATER model**		
	AIC	**df**		**AIC**	**df**		**AIC**	**df**
Variant 1	1687.341	7	Variant 1	1730.838	9	Variant 1	1711.062	7
Variant 2	1634.814	9	**Variant 2**	**1624.151**	**11**	**Variant 2**	**1636.928**	**9**
Variant 3	1660.959	9	Variant 3	1644.601	11	Variant 3	1708.461	9
Variant 4	**1626.04**	**11**	Variant 4	1633.022	13	Variant 4	1653.692	17
Variant 5	1711.546	17	Variant 5	1740.187	19			

Table 4.7 Best models' parameters (Σ $\mathcal{L}$ indicates the Likelihood sum across conditions)

Variant 4 (Pure DDM)		Variant 2 (Extended DDM)		Variant 2 (LATER model)	
AIC	1626.04	**AIC**	1624.151	**AIC**	1636.928
Σ $\mathcal{L}$	802.0199	**Σ $\mathcal{L}$**	801.075	**Σ $\mathcal{L}$**	809.4638
$a_{mFP=0.6s}$	0.464	$a_{mFP=0.6s}$	0.362	$a_{mFP=0.6s}$	0.262
$a_{mFP=1s}$	0.618	$a_{mFP=1s}$	0.575	$a_{mFP=1s}$	0.293
$a_{mFP=2.4s}$	0.812	$a_{mFP=2.4s}$	0.765	$a_{mFP=2.4s}$	0.326
$\tau_{mFP=0.6s}$	0.19	τ	0.211	$\mu_{c=0.15}$	0.82
$\tau_{mFP=1s}$	0.202	s_t	0.051	$\mu_{c=0.39}$	0.956
$\tau_{mFP=2.4s}$	0.202	$d_{c=0.15}$	3.094	$\mu_{c=0.59}$	0.979
$d_{c=0.15}$*	3.253	$d_{c=0.39}$	5.242	$\mu_{c=0.75}$	1.008
$d_{c=0.39}$	5.139	$d_{c=0.59}$	5.491	$\mu_{c=0.84}$	1.012
$d_{c=0.59}$	5.576	$d_{c=0.75}$	6.132	σ	0.152
$d_{c=0.75}$	5.823	$d_{c=0.84}$	6.739		
$d_{c=0.84}$	5.728	η	0.021		

The estimated parameters for the two DDM models are not so different, as shown in Table 4.7. The d parameter increases with increasing contrast, but the relationship is inverted for the last two drift rates with the Pure DDM, which may be due to random noise or saturation of stimulus strength. More interesting are the estimated τ parameters, which are very similar in the Pure DDM for the three mFP conditions (which were actually estimated to be identical with the last two mFP), establishing that the only relevant difference in sensory-motor component may be at most between the first and second mFP, and even in this case, the difference is very small (around 10 ms) but enough to produce a relevant change in Likelihood and AIC value. Note also how the parameter a is clearly increasing for both

* The "c" in this table stands for contrast level

versions, with estimated values not so dissimilar for one another. This clearly hints to an increase in threshold distance with increasing mFP, a finding that goes in contrast to what found previously (Jepma et al., 2012) with a Choice RT task, but it is however consistent with Posner et al., (1973), who found a speed-accuracy trade-off for different FP conditions. This will be analysed in details in the Discussion.

The Vincentized distributions of RT are shown in Figure 4.6 (the red lines refer to the Extended DDM fitting, which provided the preferred fitting). For each condition, we plotted the mean, skewness and standard deviation averaged across participants. The effect of contrast and mean FP on standard deviation was not significant (contrast: $F_{4,44} = 2.508, p = 0.055$, mFP: $F_{2,22} = 2.973, p = 0.072$), and their interaction was not significant either ($F_{8,88} = 0.381, p = 0.928$). The usual decreasing in standard deviation observed with increasing stimulus intensity was not clearly observed in this case.

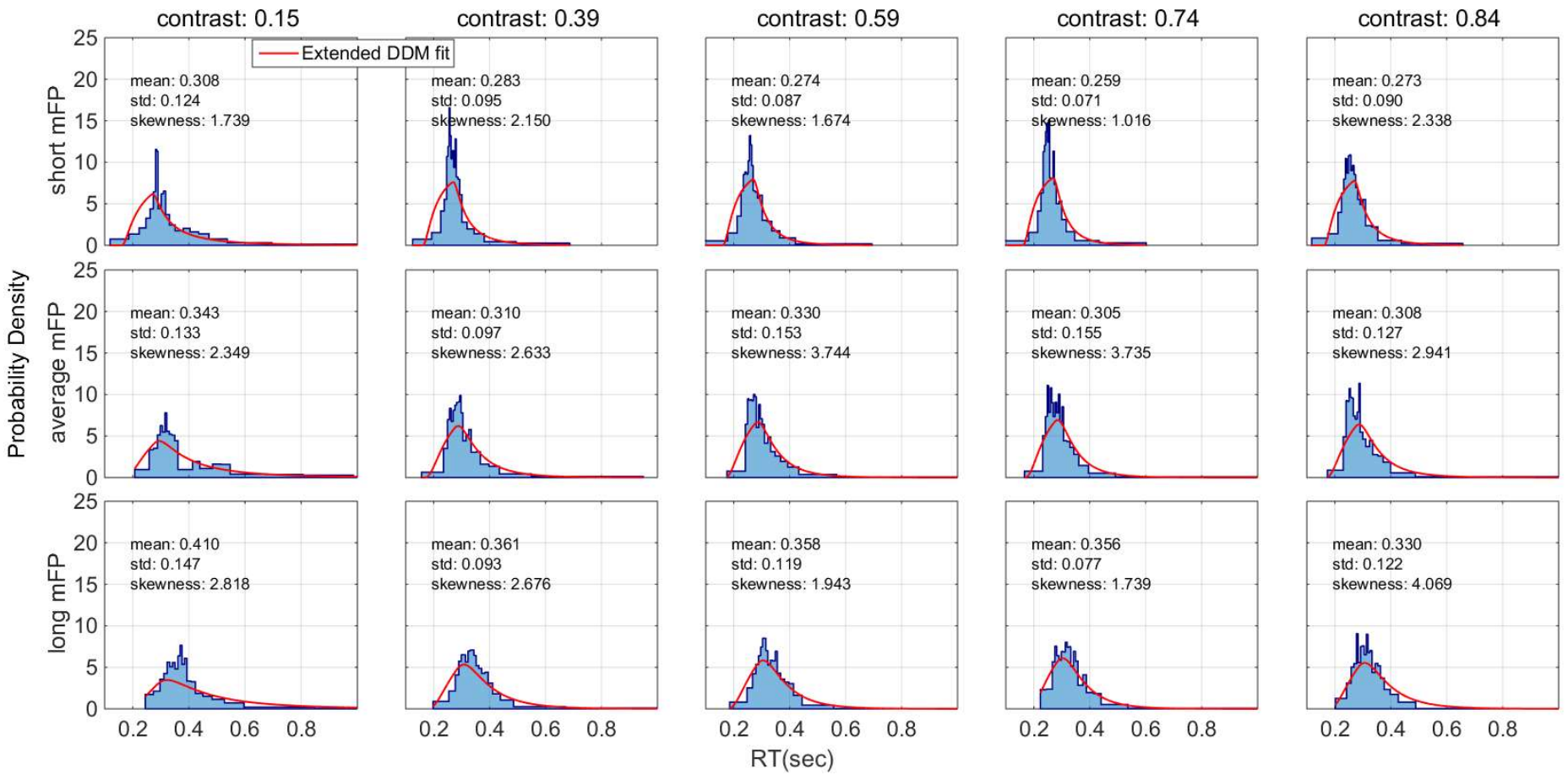


Figure 4.6

Vincentized RT distributions for each condition. The red line corresponds to the fitting of the Extended DDM, in the variant with free α conditions for each mFP condition. The other parameters are fixed for all the remaining conditions. In each panel, we indicate the average (across participants) value for the mean, standard deviation (std) and skewness of each distribution.

4.3.5 Rate Distributions

As detailed explained in Chapter 3, distributions with very fast responses will result, in the rate domain with the standardising procedure, in a distribution with a peaked shape, not resembling the original, individual rate distribution. We have seen how this study is also affected by a number of fast responses (**Section 4.3.2**) which appeared to be dependent on stimulus predictability. To solve this problem without forcing the distribution to appear Normal in the rate domain, we used the STD-IQR method (**Section 3.3**). Figure 4.7 shows both original-standardised rate distribution (blue lines) and modified ones (with the STD-IQR method, yellow lines). Note that the most affected distributions are the ones in the short mFP conditions, corresponding to those conditions with most fast responses. On the other hand, the STD-IQR method did not eliminate any value for any of the distributions with the long mFP condition, which in fact had the lowest amount of fast responses. By applying the Kolmogorov-Smirnov test (with a critical level of $\alpha = 0.01$) we revealed that 6 distributions out of 15 were not significantly different from a Normal Distribution. Four of these 6 distributions came from the long mFP condition. We also considered the effect of truncation: if a distribution is normal in the rate domain, it must be truncated at zero or near zero (Harris & Waddington, 2012). However, by inspecting individual rate distributions, we saw that all the distributions were clearly far away from zero, and therefore we did not consider truncation as a problem in our dataset.

Overall, even though the LATER model did not provide the best fitting, the distributions still appeared approximately normal in the rate domain, at least with long mFP. With medium and, more evidently, short mFP, the non-normality may be due to the fact that the distributions still contain contaminant responses, not eliminated by STD-IQR which may affect the distribution shape. The process is therefore still a mixture of a normal distribution and another component which becomes less and less active as mFP increases. Note that this component appears to be correlated with anticipations (which seems to be larger with short mFP, and decreasing with increasing mFP, as shown above), and it may be due to participant's motor preparation to stimulus onset. See Discussion for further details.

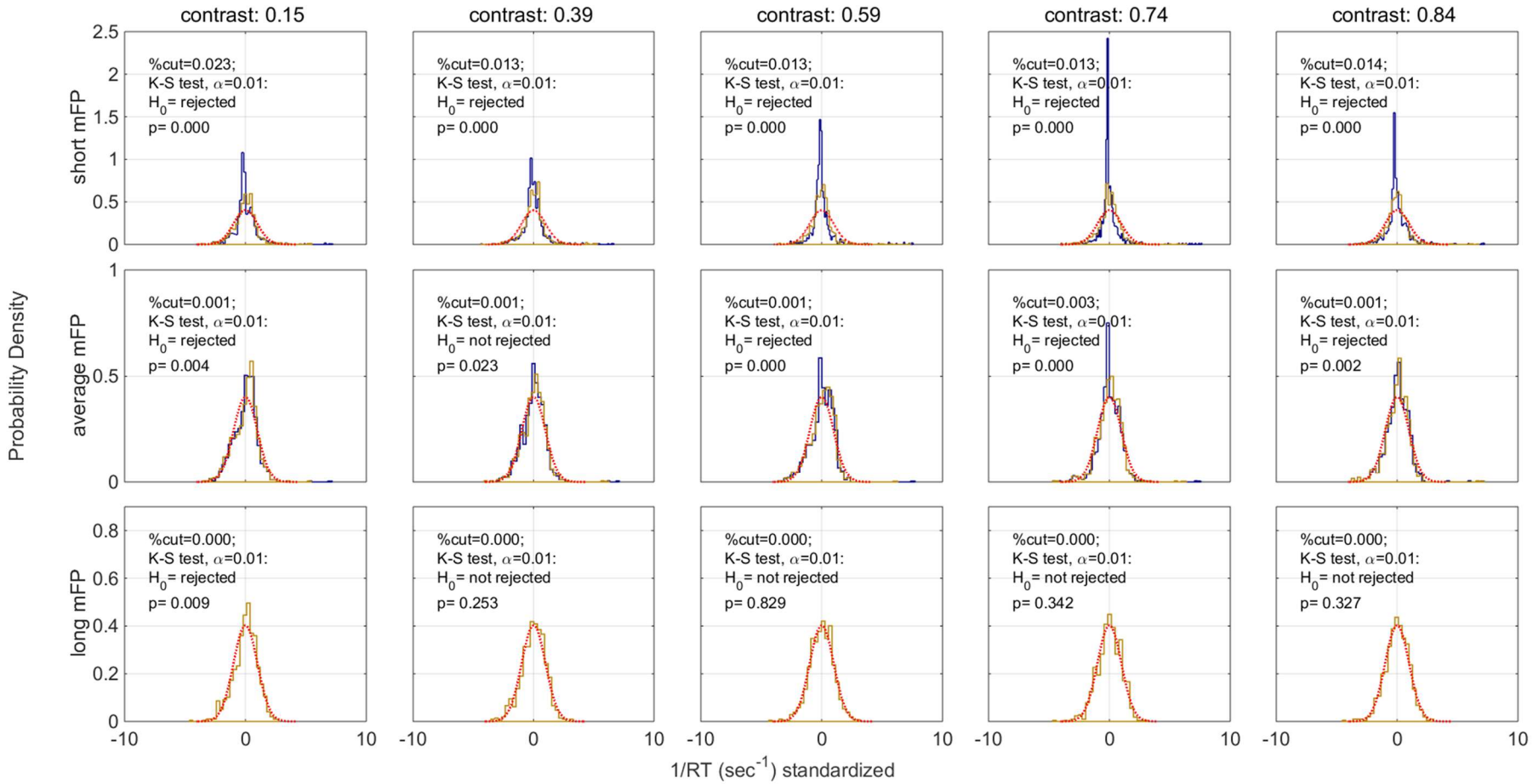


Figure 4.7

Aggregated and standardised rate distributions for each condition. The blue lines represent the original dataset, the yellow lines represent the dataset after STD-IQR method for cutting contaminant responses. Note that in the long mFP condition no contaminant is found, and the two lines overlap. In each panel there is indicated the amount of cut responses, and the result of the Kolmogorov-Smirnov test. For comparison, we plot a Normal Standard Distribution on each panel (red dashed lines).

4.4 Discussion

As seen in the introduction, most of the literature on FP and DDM indicates that by changing the FP condition we are affecting the sensory-motor parameter τ (e.g. Jepma et al., 2012; Seibold et al., 2010). We found a different and slightly more complex result: either the threshold distance a (with Pure DDM), or both threshold distance and sensory-motor component τ (with Extended DDM), appeared to increase by increasing the mFP. Note that, even when τ resulted to be affected, it was only between the first two mFP conditions, as the second and third τ estimated values were equal. One reason for the inconsistency between our results and previous studies' may be due to the fact that other studies did not take into account mixed models, but only models where either τ, d, or a were affected. As explained in the introduction, an increase in a would imply slower RT and lower accuracy, and this is precisely what we found in this study, even though the difference in accuracy (measured by using anticipations) was not significant. Interesting, Jepma et al. (2012) also found the same results, which is inconsistent with their finding that mFP affected τ (changing τ should not have any effect in accuracy), and they interpreted it as an early sampling of information.

Our finding that a is affected by mFP manipulation is particularly relevant in terms of decision strategy and optimality: if the participants are using RR_m, RR or RA (one of the decision strategies that are function of the total trial time), then by increasing the FP we expect an increase of the decision threshold which is exactly what we observed. This appears to be a strong evidence for the hypothesis that participants are adjusting their response depending on the total non-decision time in the trial. However, this is not the only explanation. With Simple RT a change in the threshold separation has been sometime interpreted as a change in the prior probability of stimulus appearance (Carpenter & Williams, 1995). For example, a may be directly connected with the predictability of the stimulus (that is, to the variance of the corresponding FP). In this case, shortest FP would correspond to lower variance, higher predictability, and thus high priors which results in low threshold separation a. Thus, observing an increase in the threshold separation parameter is a necessary but not sufficient condition to support the optimality hypothesis for RR_m, RR and RA. In the next Chapter, we will use an experiment in which

we can clearly differentiate between an increase in response time due to stimulus predictability or to increased non-decision trial time.

At face value, the Piéron's fitting appears to contradict the results from the DDM estimation. If we follow the classic definition of the γ value as the "nonsensory factor" (Scharf, 1978) or the "irreducible minimum" (Woodworth & Schlosberg, 1954), then it seems to have a direct correspondence to the τ parameter in the DDM fitting. We believe this is not the case. As the Piéron's Law fitting suggests, different mFP conditions appear to affect uniquely the γ parameter, which should mean that τ increases with increasing mFP which is, however, not what we observed (either only a or both a and τ increased, depending on the model user). This result should not invalidate the previous one: the Piéron's Law fitting is a much less powerful method to identify the stage affected by experimental conditions. Relevant to our topic of interest, it is not clear what Piéron's Law parameter should be connected with the threshold parameter (it may even affect the γ parameter itself, or a combination of parameters). Moreover, the Piéron's Law fitting procedures are typically complex (see Luce, 1986) and, with only few median data points (compared to the much more reliable distribution fitting for DDM), we think that it should be used just to confirm that contrast follows Piéron's Law, and not to infer details of the decision process' stages.

In fitting the DDM, we also wanted to compare the Extended and the Pure DDM goodness of fit for the Simple RT dataset. In fact, the Extended DDM has been recently used on Simple RT (Ratcliff & van Dongen, 2011; Patanaik, 2014; Ratcliff & Strayer, 2014), but a comparison between it and the simpler Pure DDM was not provided. This seems particularly relevant, as the Extended DDM was initially adopted in order to explain different pattern for correct and error distributions, and it is not clear whether it is really useful to apply it in situations where no incorrect distribution is obtained. We concluded, for this study, that the different is extremely small, and, overall, both models provided roughly the same results: the variability across trials for drift rate and sensory-motor component (s_t and η), which distinguishes the two models, were low. As we will see in the next Chapter, sometime the Pure DDM can even overcome the Extended DDM by providing a lower AIC value. This is interesting

as, in terms of optimality, the Pure DDM seem to be able to offer a valid and powerful account for RT and accuracy (Bogacz et al., 2010; Balci et al., 2011; Simen et al., 2009).

This experiment also provided us with some tangential information that was not strictly related with optimality, but nevertheless added useful insights in the decision making process in Simple RT.

The analysis of the relationship between FP duration and RT showed that RT decreases non-linearly (possibly exponentially, Figure 4.5) towards the mFP, in a way that matches the FP distribution itself. This shows how participants were more prepared (so to obtain shorter RT) as the cumulative distribution function increased. Note that this interesting relationship is not modelled by any of the sequential sampling models generally used in the decision making field, as they do not take into account the FP time. However, extending the model as to include the FP does not seem to be sufficient to model this phenomenon: we tried a simple model in which the starting point of the process was the beginning of the trial (and not the stimulus onset as usually done), and the accumulator consisted of white noise until the (simulated) appearance of the stimulus, whereupon the signal was noisily accumulated with a constant drift. Like our experiment, the stimulus onset time was drawn from an exponential distribution. However, this could not reproduce the FP duration and RT relationship found here. To reproduce this effect we had to add a moving boundary (by using a method similar to the one described in Drugowitsch et al., 2012), in particular, a decreasing boundary. This model would simulate a participant that willingly decreases the amount of information required before committing to a decision as time passes. Thus, this would simulate a participant that takes the FP into account. For example, during a condition with a very short FP, a participant that already waited for a long time would respond as soon as the minimum amount of information reaches his retina (as the probability that the stimulus appeared already is high). Is it conceivable that, when enough time passes, the participant would respond even though *no information* is provided, but purely based on the flow of time. With this approach, we could reproduce several patterns of the FP duration and RT relationship. It is easy to see how powerful this approach is, especially in explaining Simple RT, anticipations, relationship with prior beliefs and stimulus predictability. However, the developing of sequential sampling process with moving boundaries for

decision making is in its infancy, and a precise procedure would involve finding the optimum boundary for a set of optimality criteria by a dynamic programming approach, which is well beyond the scope of the current work. Even without developing such a complex model, we suggest the possibility that the effect of FP duration on RT may serve as a prediction for testing models in which the flow of time itself affects the response: the value of the accumulator is also function of time. Finally, we observed in this experiment, and even more remarkably in the experiment in the following chapter, a slight increase in RT after the mFP. It is possible that with very long waiting time, participants lost focus on the task, but up to now, we are not able to provide a more precise explanation.

The analysis of the distributions of responses in the Rate Domain and the fitting with the LATER model gave us important information about the Rate-Hypothesis. Note that this is the first time that such analysis in the rate domain had been performed for Simple RT task. We observed that distributions were very close to a Normal Distribution for the long mFP condition, but they got progressively less Normal with shorter mFP. This could partly be due to the effect of fast responses, which we tried to contain by using the STD-IQR method, but some responses may still have contaminated the dataset. In both the saccadic and manual RT literature, it has been hypothesized that these responses are generated by a different mechanism (Noorani and Carpenter, 2011; Noorani, 2014; Ratcliff, 1993; Ratcliff & Tuerlinickx, 2002; Ulrich & Miller, 1994), and in this work we see how this mechanism appeared to be connected to the mFP condition. Fitting the LATER model, a model which generates Normal Distribution in the rate domain, provided us with a way to compare this simple model with the more flexible DDM. From the results, it appeared that DDM better accounted for the observations, as it had a lower AIC value than the LATER model. This general result could be due to contaminant responses in the first two mFP conditions, and in fact the summed log likelihood values for the long mFP was lower for the LATER than the DDM model. The Simple RT paradigm, with its high percentage of fast responses, and with the possibility for the participants to base their response on either stimulus predictability or stimulus perception, is probably unsuited for analyses in the rate domain, and thus we delayed a complete examination of the next Chapter.

Chapter 5

Foreperiod and Reward-to-Warning Signal Interval

5.1 Introduction

In the previous Chapter, we analysed the effect of FP on RT for Simple RT in the context of optimal decisions. We used different mFP conditions as a way to manipulate the delay across trials. In this Chapter, we will use a different way to manipulate the total delay across trials, by either changing the mFP or the mean reward-to-warning-signal interval (mRWI). We will refer to the total, average delay across trials (coming from the sum of mFP and mRWI) as D (as illustrated in Figure 2.1). This time we will use a Choice RT task, to allow a more direct comparison with similar studies in the literature.

Whereas the effect of FP on Simple RT has been well documented (see previous Chapter), only few studies have analysed the same effect on Choice RT: Frowein and Sanders (1978) in a 4AFC found a clear difference in RT between two constant FP of 1.5s and 10.5s), with an increase of about 30ms and a slightly lower accuracy for the long FP condition. Vallesi, Shallice and Walsh (2006) found a similar increase in a 2AFC with 3 FP conditions (0.5s, 1s and 1.5 sec) but the accuracy obtained was not reported. Similarly, Rolke & Hofmann (2007) found reduced RT and increased accuracy following short FP. Conversely, the already mentioned work of Jepma et al. (2012) found that RT reduced by approximately 52 ms with shorter FP, (by using two constant FP: 0.3s or 2.5s), and accuracy slightly increased for the long FP condition, in the condition where accuracy was stressed. Thus, it appears that manipulating the FP exert a similar effect for both Simple and Choice tasks, at least in terms of RT, whereas the effect on accuracy is mostly undocumented or inconsistent across these few studies.

The effect of increasing the delay across trials D in terms of optimal decision has not been investigated extensively. Recall, from **Section 2.6**, that 3 of the 4 decision rules used here are function of $D_{tot} = D + \tau$ and that by increasing D_{tot} we expect an increase of both RT and accuracy (in terms of DDM, this would result in an increase in the threshold separation a). The design commonly used in these types of studies is illustrated in Figure 2.1, bottom diagram: upon participant's response, the trial automatically start after a delay (response to stimulus delay, RSI). In this design, no warning signal or FP is used, and changing the delay across trials correspond to changing the mean RSI time (mRSI). Bogacz et al. (2010) used different RSI (D=0.5s, 1s, 2,s) with fixed condition length (7 minutes in the first experiment, 4 minutes in the second). In the first experiment they used a random dot motion task, finding a significant effect on accuracy, but not on RT. In the second task (asterisk counting task) the effect on RT and accuracy was generally significant, but the difference in RT between the first and second condition was not. When participants were segregated according to their performance, the empirical RT and accuracy for the good performance group closely matched the theoretical predictions (this experiment will be discussed again in the Chapter 8, regarding quantitative optimality predictions with DDM). Simen et al. (2009) also manipulated the non-decision time by changing the RSI (D =0.5s, 1s, 2s) in a 2AFC random dot motion discrimination task. Both RT and accuracy significantly increased with increasing RSI, as predicted if participants were aiming at optimal performance. Also, by fitting the Extended DDM it was showed that the parameter affected by changing RSI was the threshold separation, as expected. In this Chapter we aim to test a similar prediction, but with a more complicated design, in which different decision components can be better distinguished.

5.1.1 Separating Stimulus Predictability and Optimization

In this Chapter we will conduct a qualitative analysis of optimality and, in particular, on the effect of the delay D and sensory-motor component τ on RT. One of the hypotheses of the previous Chapter was that increasing the FP would affect the threshold separation for the DDM. This stems from the fact that, in the context of decision rules, the threshold separation is function of D_{tot}, which is the sum of the delay across trial D and sensory-motor component τ (see Figure 2.1). We found that a increased

with increasing FP (that is, by increasing D). We also found a less strong (and inconsistent across the two versions of DDM) effect on the τ parameter. However, we mentioned the possibility that the parameter a was not actually affected because of the decision rule optimization, but because it could have been affected by the stimulus predictability. For Simple RT, in fact, the starting point of the process may as well depend on the predictability of the stimulus itself, so that if a participant knows that the stimulus will appear very soon, the threshold will be lowered. In the previous experiment, this appeared to have an effect on the threshold separation, but it may as well have an effect on the sensory-motor component τ. In the previous experiment there was no way to distinguish between whether the effect was due to stimulus predictability or to optimality adjustment due to longer trial length. To investigate this point, in this experiment we will try to distinguish between the two. We used the experimental design illustrated in Figure 2.1 (top diagram), in which the total delay is a sum of two delays in the trial, the FP and the RWI. The RWI corresponds to the time after the participant's response, and *before* the next trial starts. The warning signal at the beginning of each trial clearly indicates the separation between the end of the RWI and the next trial. If participants were estimating the probability of stimulus appearance, this estimation would start from the warning signal, not from the response at the previous trial. That is, the estimation process would be uniquely carried out during the FP time. Thus, during the RWI, there should be no "estimation process" going on, and the RWI only contributes in increasing the total trial time. We ran 3 blocks in which the mean FP (mFP) increases (but keeping an equal, constant RWI), and 3 blocks where the mRWI increases (with an equal, constant FP). To keep the design similar to the previous experiment, the FP in the constant RWI conditions was drawn from an exponential distribution (we will call this factor *variable* FP) and, similarly, the RWI in the constant FP conditions are drawn from an exponential distribution (*variable RWI* condition). Thus, all throughout this chapter, the variable RWI condition corresponds to the condition with constant FP, and the variable FP condition to the one with constant RWI. The sum of the constant component and the mean of the variable component, plus other delay in the trial due to warning signal, will be referred as delay D. Hence, we will have 3 D conditions (a more detailed explanation will be provided in the Method section).

Note that this design is clearly different from the one used in the previous studies (Section 5.1). In these works, only one delay is used (the RSI), and thus varying it would possibly affect both the stimulus predictability and the decision rule's value (due to changing in the total trial delay), making it impossible understand whether differences in RT and accuracy were due to one or the other effect.

An important difference from the previous experiment regards the experiment duration: whereas in the previous experiment we used a fixed number of trials, here we used a fixed time instead. This little modification appeared to make people more aware of the reward maximization task and has been recently used to test differences in behaviour for different RSI (see Bogacz et al., 2010; Simen et al., 2009).

5.1.2 Predictions in terms of DDM Parameters, RT and Accuracy

In fitting the DDM, we have numerous variants to choose from. In our design, where we changed the total trial delay D in two ways (by using variable FP or variable RWI), several assumptions are possible. For example, we can imagine that increasing the mFP will affect the motor-sensory component τ (due to stimulus preparation) and the threshold a (due to decision rule's optimization), whereas increasing the mRWI will only affect the latter. However, many other assumptions are possible. We selected 9 different variants (valid for both the Extended DDM and the Pure DDM) which appeared to capture the most diverse effects that such design may have on the DDM parameters. In all these models, we assume that the drift rate d is constant for all the conditions, as the stimulus is kept the same throughout the whole experiment. In order to summarize these variants, we use a matrix representation (Table 5.1 and Figure 5.1). The Table 5.1 presents the organization of the matrix representation: each column represents a different D condition (from left to right, $D = 1.6s, 2s, 3.4s$), and the two rows represent either the constant RWI (variable FP, top row) or the constant FP (variable RWI, bottom row) condition. Note that each D condition is a sum of the constant delay plus the mean of the variable delay plus $0.5s$ delay for the warning signal. Figure 5.1 shows the associations of the parameters with the 9 variants used. As before, different columns correspond to different D condition, and the two rows correspond to variable FP and variable RWI respectively. In each matrix within the figure, when a

parameter is shown within a cell, it means that a parameter is estimated uniquely for that cell. When the parameter is shown on top of a column, the parameter is assumed to be the same for all the cells in that column (that is, for the whole D condition corresponding to that column), and similarly when it is placed on the side of a row, it means that it will be the same for all the cells in that row (that is, for the whole variable FP or variable RWI condition). For example, in Model 6 we assumed a different τ value for each D condition in the variable FP, but the same τ value for every condition of the variable RWI. Plus, we assumed a different a value for each condition. This would result in a model with a total of 11 parameters (4 τ, 6 a and 1 d).

We will start by describing the models in which the threshold separation a is not taken into account, and different conditions are deemed to affect only the τ parameter. This may happen because participants are not optimizing based on D_{tot}, and the τ parameter varies depending on other reasons, for example motor preparation. We chose three variants which capture most of the scenarios. Model 1 shows a version in which participants sensory-motor response depends uniquely on the D factor, irrespectively of D increasing because of the FP or the RWI. Conversely, in Model 2, τ depends uniquely on the variable FP-RWI factor, irrespectively of the actual length of the trial D. Model 3 presents a more intuitive possibility, in which τ is function of the predictability of the FP (that is, it increases with increase variability of the FP). Thus, different τ are estimated for increasing mFP, and another τ is estimated for the constant FP used in the variable RWI condition. For all of these models, RT should increase with increasing τ, but accuracy should not be affected.

If participants are maximizing a decision rule that is function of D_{tot} (equal to $D + \tau$), then the threshold separation a must change depending on both D and τ. This is implemented in Models 4, 5 and 6, in which a different a is used whether D, τ or any combination of those are changed. For example, in Model 4 a is assumed to vary for different D conditions only, as τ is also varied for different D condition, and no other a parameters need to be assumed. Instead, for Models 5 and 6, we need to estimate a different a for each cell, as either D or τ are varying. For any of these models, RT and accuracy should increase when a increases.

Finally, we use variants that stem from the previous experiment, in which we observed that longer FP affected a, but only weakly τ. To test for this hypothesis, we use three more Models: 7, 8 and 9, which mimics the Model 1-3 for τ, but with the parameter a instead. For these models, RT and accuracy should increase with increasing a.

Table 5.1 Organization of the study (D is the sum of constant delay, variable delay, and warning signal delay)

	$\boldsymbol{D = 1.6s}$	$\boldsymbol{D = 2s}$	$\boldsymbol{D = 3.4s}$
$\boldsymbol{CONSTANT\ RWI = 0.5s}$	$mFP = 0.6s$	$mFP = 1s$	$mFP = 2.4s$
$\boldsymbol{CONSTANT\ FP = 0.5s}$	$mRWI = 0.6s$	$mRWI = 1s$	$mRWI = 2.4s$

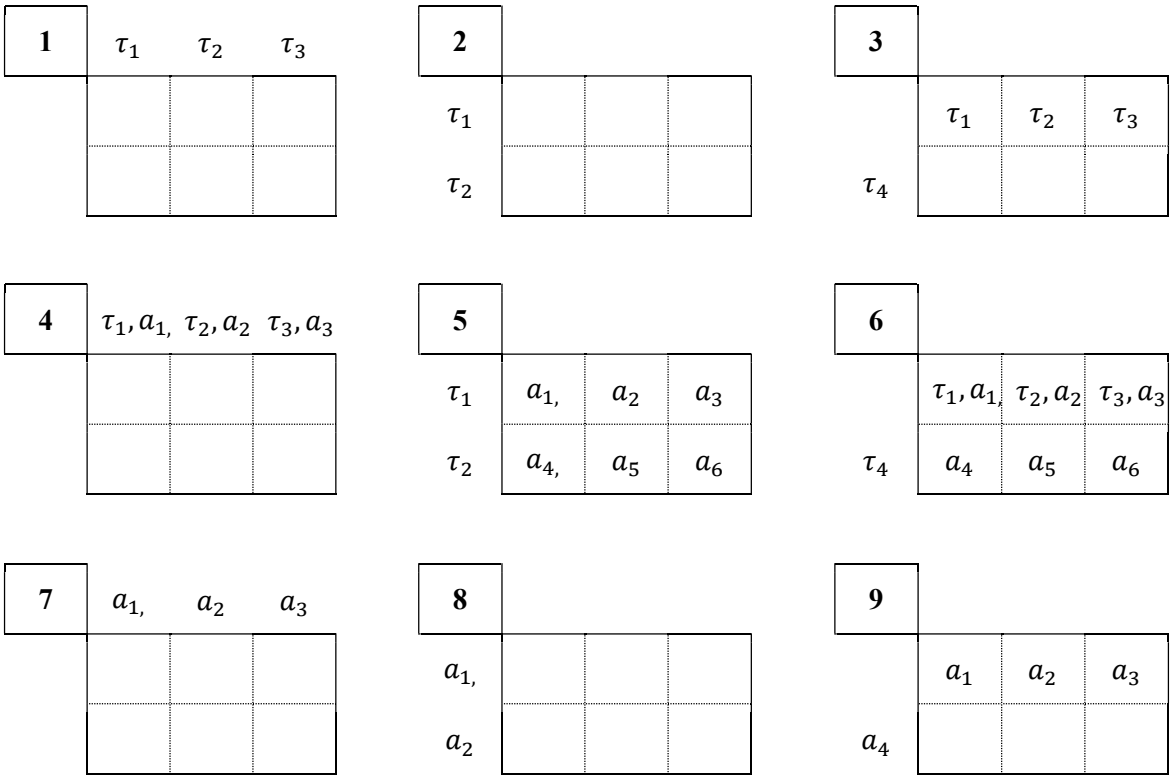


Figure 5.1

Illustration of the 9 different variants used for fitting the Pure DDM and the Extended DDM. Each cell indicates one experimental condition: the columns correspond to different ***D*** conditions, and the row to variable FP/variable RWI conditions (see Table 5.1). When a parameter is indicated on the side of a row or on top of a column, it indicates that that parameter is fixed across the whole row/column condition. When a parameter is within a cell, it indicates that the parameter is fixed within that condition, but varies on the other conditions. The drift rate parameter is always considered fixed across all conditions.

Extended DDM and Pure DDM will be tested an all the 9 variants, whereas the LATER model will be tested on the 7, 8, and 9 variants (as it does not have the τ parameter).

5.1.3 Aim of Experiment 5.1

1) The main aim of this experiment is to investigate the effect of different mFP and mRWI on accuracy, RT, and parameters of the diffusion model so that we can understand whether participants' response is affected by the delay across trial, as predicted by several decision rules, and whether this effect is related to the stimulus predictability or to the effect on the value of the decision rule.

2) In the previous study we observed a peculiar relationship between the FP duration and RT (Figure 4.5). Here we analyse the same relationship in the context of Choice RT. The comparison between FP and RWI duration effect will allow us to check if the decreasing relationship found for FP reflects participants' stimulus predictability estimation or a modification due to a decision rule affected by delay across trials.

3) As before, we are interested in comparing 3 decision models: Pure DDM, Extended DDM, and LATER model. We are especially interested in the prediction of the DDM models, as they can give a precise indication with regards to what parameters changes with different conditions.

4) Similar to the previous Chapter, we analyse the normality of distribution in the rate domain. For manual Choice RT, this has been done only once (Harris et al., 2014), in a remarkably different design.

EXPERIMENT 5.1

5.2 Methods

Participants. 14 participants (9 males and 5 females) took part in the experiment. All participants had normal or corrected-to-normal vision and no known neurological conditions.

Materials. The testing was conducted under constant levels of illumination. The stimuli were presented binocularly on a computer monitor (Sony GDM-F520) using the software E-Prime (version 2.0.10). Participants were positioned at 63 cm from the monitor. The stimuli were Gabor patches

(generated with MATLAB) presented in the middle of the screen. Each patch is a vertical sine-wave grating multiplied by a Gaussian windows function ($\sigma = 0.08°$), trimmed for values lower than 0.005. The spatial frequency was 4 c/deg and a contrast level was 0.2. The stimuli were presented within a grey square window (10 cd/m^2) of 5 degrees/cm. The warning signal was a cross (3.5 degrees/cm) in the centre of the screen which appeared for 0.5s. The rest of the monitor screen was kept dark (~0 cd/m^2). We used 3 variable-FP blocks and 3 variable-RWI blocks. In the variable FP block, the FP duration was drawn from an exponential distribution truncated at 10s with mean equal to 0.6s, 1s or 2.4s, depending on the block. In these blocks, the RWI was kept constant at 0.5s. Similarly, in the variable RWI condition the RWI was composed in the same way, and the FP was kept constant at 0.5s. The order of block presentation was randomized. Therefore, we obtained a total of 2x3 conditions: the first factor is the variable FP and variable RWI, the second is the average delay length of the trial (D), in seconds, which corresponds to the sum of mRWI, mFP, and delay due to the warning signal ($0.5s$). For example, with variable FP (or variable RWI) with mean 1s, the RWI (or the FP) was kept constant with length 0.5s, and thus D was equal to 0.5s (warning signal duration) + 0.6s (mean FP duration) + 0.5s (RWI duration) = 1.6 seconds. Similarly, the second D condition length was 2seconds, and the third was 3.4seconds . Note that in this experiment we did not use the minimum waiting time 0.4s (in the previous experiment this quantity was added to the variable part of the FP). This is done to avoid anticipations, as observed in the short mFP condition for the previous experiment.

Procedure. The experiment is presented as a point game. Participants were explicitly told that they needed not to be as accurate or as fast as possible, but they needed to earn as many points as possible. Each participant was tested in each one of the six blocks. The order of each block was randomized, and each block lasted 3 minutes regardless of participant's speed of responses. The trial started with the presentation of the warning signal, which disappeared after 0.5s, followed by a black dot, used as a fixation point. After the FP (variable or constant depending on the condition) the dot disappeared and a vertical line appeared within the grey window. At the same time the Gabor patch appeared on the left or on the right of the vertical line. In this way, the participants knew when the stimulus appeared and

had only to detect its location. Participants were asked to press the Z key for left and the M key for right. Upon their response, the vertical line and the Gabor patch disappeared and, after a constant or variable RWI time (depending on the condition), the next trial began. A correct response was followed by a positive feedback sound, whereas an incorrect sound was not followed by any sound. Participants earned 1 point for a correct response, and lost 1 point for an incorrect one. At the beginning of each condition, participants were given 25 points. At the end of each condition, participants were informed about the total score for that condition. To increase motivation, participants were informed that a £10 prize would be awarded to the participant that earned the most. Figure 5.2 shows a schematic representation of the flow of the trial, and Table 5.1 (above) summarizes the different conditions used in this study.

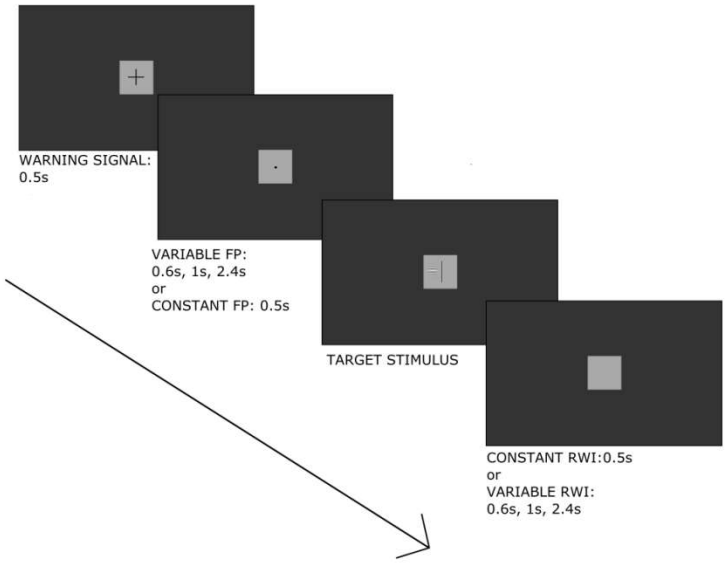


Figure 5.2

Illustration of the flow of the task.

5.3 Results

5.3.1 Behavioural Responses

The percentage of error responses was low (average across condition=2%, SE=0.8%). These responses were excluded from the main analysis and, as before, they will be analysed separately (**Section 5.3.2**). The results in terms of median of median RT (mdRT) are shown in Figure 5.3. We ran

a 2-way repeated measure ANOVA and found that both manipulating the variableFP-RWI condition and the D condition significantly affected the median rate (for the variableFP-RWI factor: $F_{1,13} = 19.047, p < 0.001$, for the D factor: $F_{2,26} = 9.334, p < 0.001$); the interaction was also significant ($F_{2,26} = 6.584, p < 0.005$). However, from the figure, it appeared that the delay had an effect only for the variable FP condition, which was confirmed by the analysis of variable FP and variable RWI conditions separately. The effect was significant for the variable FP condition ($F_{2,26} = 19.341, p < 0.0001$). Post-Hoc analyses revealed that there was a significant difference for short and long FP conditions, and medium and long FP conditions (both at $p < 0.0001$), but not between short and medium FP conditions ($p = 0.5$). As suspected, the effect of D in the variable RWI condition was not significant ($F_{2,26} = 1.305, p = 0.288$). We also segregated the participants in three different groups according to their performance (**Section 3.1**), but this did not show any obvious pattern, apart for slightly faster responses for the good performance group. The effect of D within variable RWI on mdRT was not significant in any of these sub-groups, whereas the effect of D within variable FP was significant for all of them (at $p < 0.001$). This increase in mdRT for different delays only in the variable FP condition suggests that participants were actually affected by the stimulus predictability, and not by the general increase of delay across trial, contrarily to what the previous experiment suggested. In fact, if participants were affected by the total delay D in the trial, we would expect an increase for both variable FP and variable RWI with increasing D conditions.

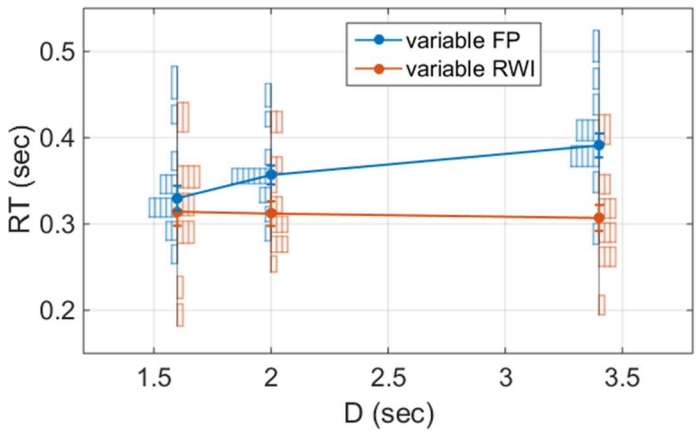


Figure 5.3

Each circle represents the median of median (mdRT) across participants, for each condition. The lines connect the 3 D conditions for each variable FP/variable RWI conditions. The vertical blocks represent the frequency distributions of the median RT for each participant. Increasing the duration of variable FP seems to be really effective in increasing the mdRT, whereas increasing the mean duration of the variable FP does not seem to have any significant effect.

5.3.2 Fast Responses and Error Responses

As before, we analysed the number of fast responses (responses shorter than 0.2s). The results are shown in Figure 5.4, left panel. Note how, for variable FP, the number of fast responses was much small than in the previous Simple RT experiment. We believe this may be partly due to the fact that the Choice RT design used here did not encourage anticipations, and that we eliminated the 0.4s of minimum waiting time from the variable FP-RWI duration, and is therefore much more difficult to predict the occurrence of the stimulus. The increase of fast responses for variable RWI condition is expected, as in this condition the FP was constant and thus the stimulus onset is much easier to predict. The number of errors were averaged across participants and shown in Figure 5.4, right panel. There seems to be a slightly lower percentage of error for the variable RWI conditions. However, none of the comparison resulted to be significant.

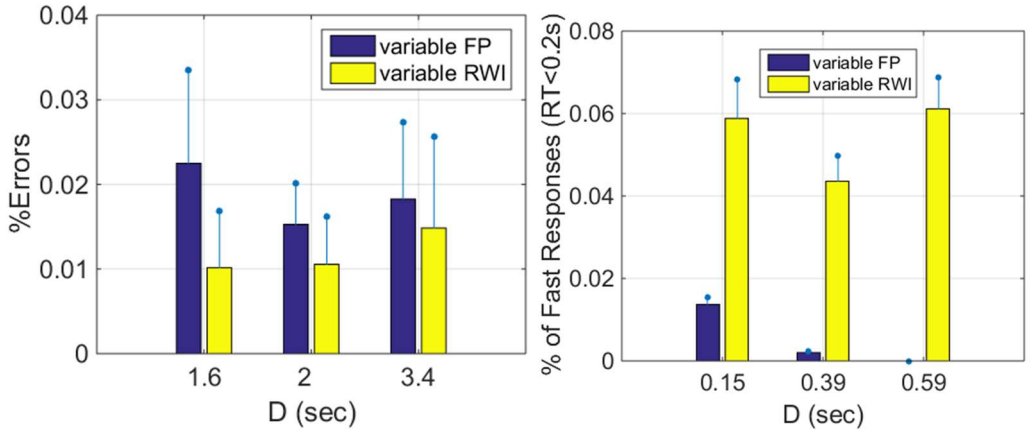


Figure 5.4

(Left panel) Percentage of error responses for each condition, averaged across all participants. **(Right panel)** Percentage of fast responses (responses faster than 0.2s), averaged across all participants. The blue bar indicates one standard error.

From the analysis of these behavioural responses, it appeared that increasing the average trial time D had, by itself, no relevant effect, as there is basically no change in RT with increasing mRWI.

Remember that the decision strategies RR, RR_m and RA are all affected by the total non-decision trial time D, and the effect should affect the percentage of errors as well. Participants however did not seem to be sensitive to this factor, which may lead us to conclude that participants were not using any of the aforementioned decision strategies. The change in RT when increasing variable FP may be explained by a decrease in the motor preparation, or by a difference in the starting point of the accumulated information (Jepma et al., 2012). By fitting the DDM, we will be able to give a more detailed account for these results.

5.3.3 Model Fitting

The Vincentized distributions with the mean, standard deviation and skewness averaged across all the participants are shown in Figure 5.5. We found a significant effect of variable FP\RWI condition (but not of D) on the standard deviation ($F_{1,13} = 13.108, p = 0.003$). Neither of the two conditions had any effect on the skewness of the distributions.

We fit the 9 models presented in Table 5.2 with the Pure DDM and Extended DDM. We also fit the LATER model. As this latter does not have the τ parameter, we used the last three variants (Model 7,8 and 9) shown in Table 5.2. The best fitting model, for both Pure and Extended DDM, was Model 3: this is the variant with different τ depending on the predictability (variability) of the FP. This model estimated four different τ values: three for the three exponentially distributed FP in the variable FP condition, and one for the constant FP in the variable RWI condition. The second best model, with a very similar AIC value, was the one with different a for different FP predictability (Model 9). This is equivalent with the previous model, but with a instead of τ. Note how, in this case, the Pure DDM appeared to generate a preferable fitting (lower AIC) than the more complicated Extended DDM. By analysing Likelihood values for individual conditions, it is clear that the Extended DDM fits the data slightly better, but the higher number of parameters made this model less preferable compared to the Pure DDM. It is possible that the Pure DDM was favoured for the Choice RT due to the lack of error distributions, which may have been only properly fitted by the Extended DDM (Ratcliff & McKoon, 2009). The difference appeared to be irrelevant for this particular experiment, as both models seem to

favour the same variants, that is to say that they point to the same underlying explanation. The Pure DDM fitting is shown on top of the distributions in Figure 5.5. AIC value for each variant are shown in Table 5.2, and estimated parameters for the best variants are shown in Table 5.3. In this Table, the parameter index is referring to the corresponding index in Table 5.2, so that, for example, τ_2 for Model 3 refers to the value for condition with $D = 2s$ and variable FP. Note how the estimated parameters for both Pure and Extended DDM strongly confirmed the idea that τ depends on the predictability of the FP, as their value closely follow the FP variability. In fact, τ increases with increasing mFP (which correspond to an increase in variance) during the variable FP condition. In the variable RWI, the unique estimated τ is lower than the lower τ in the variable FP condition, as in this case the FP is constant.

The LATER Model, used as a way to test the Rate-Hypothesis and compare it with the DDM, appear to be able to reasonably account for the data. The best Model, in this case, is the same as the second preferred model for the DDMs. Note that the LATER model does not have a parameter τ to model for sensory-motor delay.

Table 5.2 Fitting results for DDM and LATER model

Pure DDM			**Extended DDM**			**LATER model**		
	AIC	**df**		**AIC**	**df**		**AIC**	**df**
Variant 1	670.7935	5	Variant 1	676.5152	8	Variant 7	670.9422	5
Variant 2	650.7415	4	Variant 2	656.7119	7	Variant 8	651.5873	4
Variant 3	**647.7999**	**6**	**Variant 3**	**653.8526**	**9**	**Variant 9**	**649.4133**	**6**
Variant 4	673.0621	7	Variant 4	679.0051	10			
Variant 5	652.1912	9	Variant 5	658.2287	12			
Variant 6	656.0217	11	Variant 6	663.3461	14			
Variant 7	670.5116	5	Variant 7	676.1423	8			
Variant 8	650.636	4	Variant 8	656.6487	7			
Variant 9	648.6185	6	Variant 9	654.6275	9			

Table 5.3 Best models' parameters

Variant 3 (Pure DDM)		Variant 3 (Extended DDM)		Variant 9 (LATER model)	
AIC	647.8	**AIC**	654.628	**AIC**	649.413
$\Sigma\,\mathcal{L}$	317.9	$\Sigma\,\mathcal{L}$	318.313	$\Sigma\,\mathcal{L}$	318.7066
a	1.746	a	1.64	a_1	0.292
τ_1	0.208	τ_1	0.219	a_2	0.309
τ_2	0.227	τ_2	0.239	a_3	0.335
τ_3	0.25	τ_3	0.26	a_4	0.267
τ_4	0.18	τ_4	0.19	μ	0.84
d	5.761	s_t	0.012	σ	0.141
		s_z	0.058		
		d	5.646		
		η	0.106		

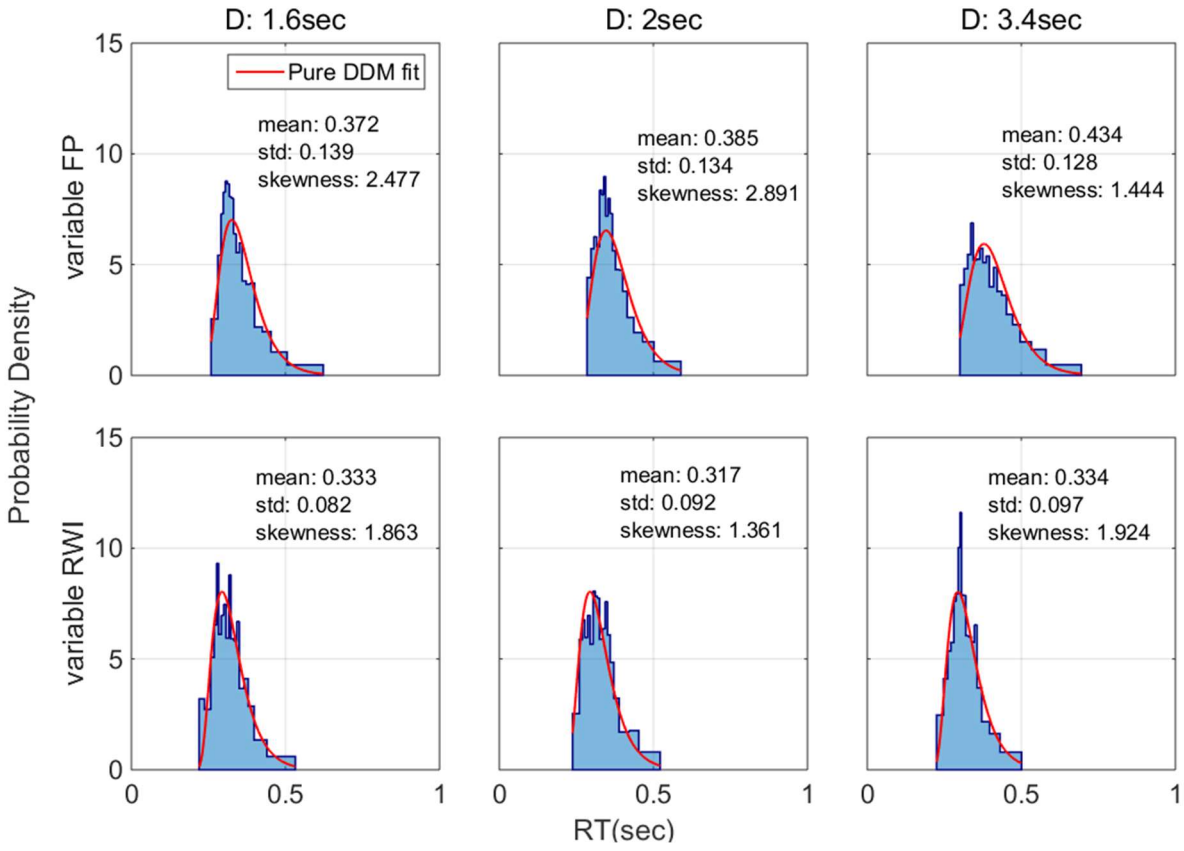


Figure 5.5

Vincentized RT distributions. The red line corresponds to the fitting model, which in this case was the Pure DDM with free τ only along different D conditions for the variable FP condition. The different sensory-motor component (τ) for the top panels corresponds to shifting the estimated distribution on the horizontal axis, leaving the shape unchanged. Even though the shapes in the top panels' distributions do indeed vary for different D conditions, the fitting procedure found the model with free τ *and* free α to less parsimonious (higher AIC) than the fitting with only free τ, which was then preferred.

5.3.4 Foreperiod Duration and RT

As in the previous experiment, we analysed the relationship between FP-RWI duration and the RT (see **Section 4.2**). In this case, however, the plot of FP-RWI duration starts from 0 instead that from 0.4s (as 0.4s was a constant time added to the random FP in the previous study). We obtained a very similar result as in the previous experiment, with a sharp decrease for the shorter FP, but with no decrease for variable RWI (see Figure 5.6). As before, the fitting of an exponential function was clearly poor (average R^2 for variable FP was lower than 0.5, fitting not plotted). By inspecting the data, the strength of the decrease in the variable FP seemed to be strongly dependent on the D conditions, with high D having an almost flat relationship. The general findings was similar to what found in the previous experiments: participants appeared to base their response on the variable FP, probably trying to estimate $P(T < t)$. This would have no effect for the variable RWI condition, where the FP is maintained constant. The slight increase we observed for the variable FP condition, after the mean FP in time, is present here as well.

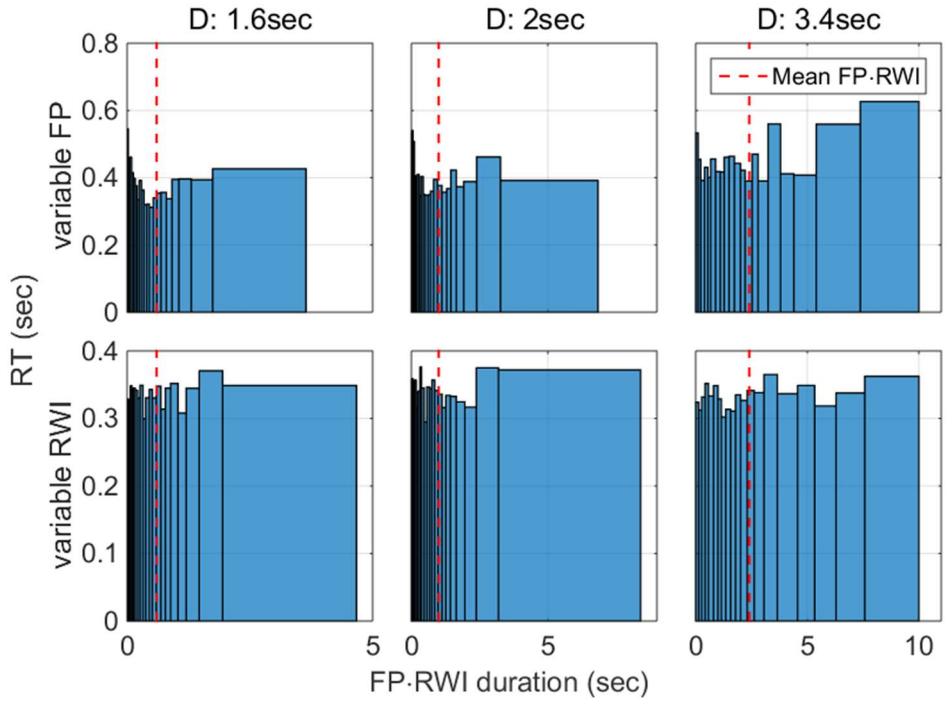


Figure 5.6

Relationship between FP duration and RT (three top panels) and RWI and RT (three bottom panels) for each D condition. The red-dashed line indicates the mean FP (top panels) or mean RWI (bottom panels) for that condition. Note how the rightmost panels are differently scaled on the horizontal axes. We can see a decreasing relationship for the variable FP condition, but no clear relationship for the variable RWI. This is expected, if we consider that participant are basing their response on stimulus expectation (recall that for variable RWI the FP is constant).

5.3.5 Rate Distributions

As in the previous experiment, the rate distributions were analysed by applying the STD-IQR method so to eliminate very fast responses. Figure 5.7 shows the original dataset and the one (with STD-IQR), after standardising and aggregating across all participants. The method had the largest effect on the variable RWI conditions. This is to be expected: as we have observed in the previous Simple RT experiment, as the FP variability decreases, the number of fast responses increases, and here the FP is constant (no FP variability) in the variable RWI conditions. The original distributions presented the classic peak shaped and, when the fast responses are eliminated, they appear remarkably closer to a Normal Distribution. Note that for the D=3.4s condition in the variable FP case, the STD-IQR method did not exclude any point. The Kolmogorov-Smirnov test showed that the null hypothesis for a standard normal distribution cannot be rejected (with $\alpha = 0.01$) for any of the condition with variable RWI and for the longer mFP duration in the variable FP condition, meaning that the standardised rate distributions

for these conditions were approximately normal. This topic will be discussed more thoroughly in the Discussion (**Section 5.4**).

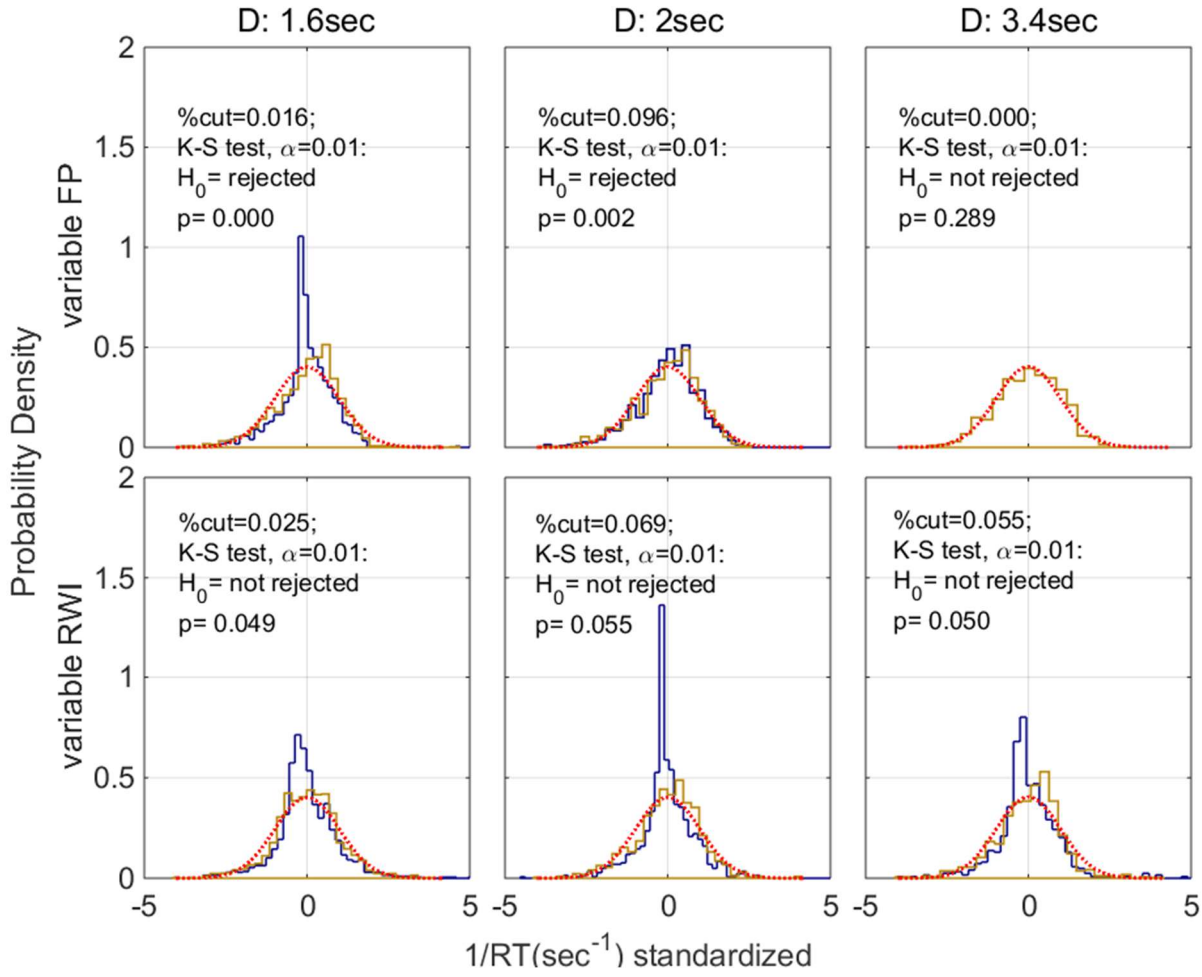


Figure 5.7

Aggregated and standardised rate distributions for each condition. The blue lines represent the original dataset, the yellow lines represent the dataset after STD-IQR method for cutting contaminant responses. In each panel, there is indicated the amount of cut responses, and the result of the Kolmogorov-Smirnov test. For comparison, we plot a Normal Standard Distribution on each panel (red dashed lines). The amount of contaminants appear to decrease with increasing D conditions, but only for the variable FP condition. No contaminant were found for the $D = 3.4s$ in the variable FP condition. For the variable RWI condition (constant FP) the amount of contaminants did not seem to be affected by the different RWI duration.

5.4 Discussion

5.4.1 Differences between Simple and Choice RT

The experiment presented in this Chapter was designed to clarify the relationship between increasing the non-decision time of the trial D and the response obtained. This parameter is particularly

relevant because is shared by three decision rules (RR_m, RR and RA) and it has a simple relationship with responses and accuracy (increasing D corresponds to an increase in response and accuracy). The experiment from the previous study suggested that participants were optimizing one of those decision strategies, as the best DDM was the one with free threshold parameter a, and it also corresponded to an increase in response time and a (non-significant) increase in accuracy. However, manipulating the FP means manipulating the predictability of the stimulus as well as total non-decision time, and we designed the current experiment to better distinguish between the two.

In this experiment, we replicated the result found in the earlier study about the increasing effect of mFP on RT, but the effect completely disappeared with variable RWI. As before, increasing the mFP in the variable FP condition corresponded to an increase of RT. Unfortunately, the very small number of errors made it impossible to capture the relationship between increasing FP and accuracy. The estimated τ parameters for both models confirmed that participants based their response on the FP variability, and not on the total delay across trials. In the previous experiment, however, we found that higher mFP affected the parameter a instead of τ. What is the reason for this difference? We propose two explanations

1) In Simple RT and Choice RT, the predictability of the stimulus affects different components, as the two tasks are actually based on two different types of decision. The main difference is that, at the starting point of the DDM accumulation for Simple RT (the stimulus onset), the participants have already had access to some information based on how much time has passed in the trial. In the Choice RT case, the stimulus onset correctly represents the beginning of the perceptual process itself, as no information can be used to make a decision *before* the stimulus appeared. Thus, in Simple RT, the threshold separation may be conveniently varied by the participants based on the prior probability, depending thus on FP variability (Carpenter & Williams, 1995). On the other hand, in Choice RT tasks, having waited for long does not give information about which stimulus is correct, and thus there is no reason to adjust the threshold separation. However, in this case, it is possible that increasing the

variability increases participant's preparation to respond, and thus affects the τ parameter. The same effect in Simple RT, if present at all, would be masked by the large effect of threshold adjustment.

2) Note how the AIC value is almost identical for the version with free a and free τ, both depending on FP variability. As estimation process are noisy and may depend on peculiarity of the dataset, when two AIC are so close is very difficult to strongly favour one instead of another. It is thus possible that different stimulus variability affects the a parameter, instead of the τ as usually assumed. One way to clarify this point would have been by employing an experiment that produced more errors, so that the effect of variability and delay may have been analysed (recall that increasing a corresponds in an increase on accuracy, whereas increasing τ does not).

5.4.2 Delay and Optimality

In **Section 2.7** we identified 3 levels of analysis for investigating decision strategies. In this and previous experiment we operated, at the same time, on the first two levels, by investigating the effect on RT and accuracy, and by investigating the effects in terms of DDM parameters. Even though with limitation, this approach can give us some interesting insights about decision rules in RT tasks. The main point we learnt in these two experiments is that participants did not take into account the total delay across trials. This finding may seem to exclude the RR, RR_m and RA decision rules, which are function of D_{tot} (which, in turn, is equal to the sum of D and τ). This result conflicts with the previous literature. As seen in the introduction, both Bogacz et al. (2010) then Simen et al. (2009) found an increasing in RT, accuracy and a parameter in Choice RT by increasing the RSI. By using the present paradigm, we can actually explain this inconsistency: these two works, in fact, did not distinguish between a delay that affects temporal preparation and a delay that does not. In practice, their RSI is similar to the FP in that it will affect more than just the delay across trials, but temporal preparation due to stimulus predictability as well. This point was not considered in the aforementioned studies. The researchers observed that a was affected by increasing delay, and concluded that this was a sign that participants were employing a decision rule that is function of delay. However, the threshold separation may be affected by variability/temporal preparation. Observing an increase in a would thus be a

misleading proof that a decision-strategy adjustment is taking place. With the design proposed in this Chapter, we provided a much more insightful interpretation of the mechanism employed by different participants.

Does the fact that participants did not use the delay across trial mean that RR, RR_m and RA are decisively excluded by the pool of possible decision rules used, and we should thus only focus on the BR decision rule? We believe that this is not the case. As we will see in the following Chapters, both RR_m and RA provide an excellent fit in some experimental conditions, and they may be several reasons why participants did not seem to be sensible to delay modification (finding confirmed in the following experiments). We will touch this topic again in Chapter 7, Experiment 7.1.

5.4.3 FP-RWI Duration and RT

By comparing this experiment with the previous one, we learned that the Simple RT and Choice RT have several differences, but also some similarities. Firstly, in both paradigms, there was a tendency to produce faster RT with longer FP (we are not referring to the mFP, which produces *slower* RT, but to the real FP duration within each mFP condition). This seemed to be uniquely affected by the FP condition, as there was no difference for the variable RWI task (that is, different D conditions but constant FP). Since this phenomenon took place for both Simple and Choice RT, it may be due to the temporal preparation of the participant to the stimulus onset, which depends on the probability that the stimulus appeared already in the past: $p(T < t)$. If this were true, we expect the relationship to be flatter with longer FP, and in fact this is what we observed by inspecting the dataset: for both Simple RT and Choice RT, there is an evident difference between $D = 3.4s$ and the two smaller values (the difference between D=1.6sec and D=2sec is however less clear, see Figure 5.6). Unfortunately, our method for measuring the strength of this relationship (by fitting an exponential distribution) did not seem adequate, as the R^2 resulted to be very low and the fitting unreliable.

5.4.4 Normality in the Rate Domain: DDM Simulations

As in the previous experiment, we analysed the distributions in the Rate Domain in order to check whether they were approximately Normal. Original distributions presented a peak similar to the one observed for Simple RT, which appeared to be related to FP predictability: the peak is present for all the three variable RWI conditions (with constant FP) and it clearly decreases for the variable FP conditions. In the conditions with long mFP, the distribution does not present any fast response (no value was cut from the STD-IQR method) and it is clearly near Normal (Figure 5.7). The same results were obtained for the previous experiment, in which longer (less predictable) FP generated approximately Normal distributions. The LATER model, in this case, fitted approximately as good as the DDM model, in spite of the lack of the τ parameter. It appears, in fact, that the initial sensory-motor relationship may be as well modelled by the threshold separation (a) parameter instead. The fitting of the LATER model in the previous Chapter was however remarkably poorer than for the DDM model. The results from these two studies can only be speculative: the fact that distributions appeared approximately Normal in the rate domain cannot be used as a definitive answer to the hypothesis that the brain computes the rate of response and then translate it in time (Rate-Hypothesis, **Section 2.4**). The main reason is that it may be very difficult to distinguish between a Normal distribution and a rate distribution generate by the DDM. It is believed (Ratcliff & Reddi, 2001) that a DDM can generate a "reasonably" Normal Distribution in the rate domain, but a deep investigation has not been performed yet. Of course, the distribution will mathematically not be Normal in the rate domain, but it is possible that with small sample a Kolmogorov Smirnov test may not exclude Normality. This is exactly what we tested. We ran a set of simulations with the Pure DDM and Extended DDM, for both the one-boundary and two-boundary version (Simple and Choice RT). We simulated 15 participants and, in turn, varied each one of the parameters of the model: a, τ and d for the Pure DDM; a, τ, d, η, s_z and s_t for the Extended DDM; but we did not use s_z (the variability across trials for the starting point), for the one-boundary version of Extended DDM, to replicate what normally done in the literature. When not changed, the parameters for each participant were drawn from a normal distribution (truncated at 0) which means that they were equal to the values shown in Table 5.4, with standard deviation equal to 0.1 (this was done to simulate individual variations). The mean values used for each participant were

similar to the one found in the literature and in the present experiments. We run the simulation for different samples ($N = 20, 50, 100$ and 200) corresponding to the number of simulated trial each participant performed in each condition. Each one of this simulation was repeated 30 times.

Table 5.4 Values used in the DDM simulations

Mean value parameters for the Pure/Extended DDM		Range
$\boldsymbol{a}$	1.5	[0.2 … 5]
$\boldsymbol{\tau}$	0.1	[0 … 0.4]
$\boldsymbol{d}$	2	[2 … 10]
$\boldsymbol{\eta}$	0.05	[0 … 0.2]
$\boldsymbol{s_z}$	0.05	[0 … 0.2]
$\boldsymbol{s_t}$	0.05	[0 … 0.2]

After we generated the results, we run the STDI-QR method, standardised and aggregated the distributions, and we ran the Kolmogorov-Smirnov test. We use the resulting p-value as a way to test distribution normality: higher value means that normality could not be excluded. Pure/Extended DDM for both Simple and Choice RT gave the same outcome: for $N = 20$ and $N = 50$, for the "reasonable parameter range" (that is, parameters usually found in the literature), the distributions appeared to be approximately Normal, so that a Kolmogorov-Smirnov test would not have rejected normality. For example, Figure 5.8 shows the results for the parameters of the two-boundary version of the Pure DDM. The results for the other parameters of both Pure DDM and Extended DDM are similar, both for the one and two boundaries versions. Figure 5.9 shows an example of the distribution shape in the rate domain for a certain parameter combination, in a two-boundary Pure DDM, for different sample size. A standard Normal Distribution is plotted on top. The distributions are generally slightly skewed on the right side, but the Kolmogorov-Smirnov test does not detect it when the sample size is too small, so that normality was not excluded in those cases. It seemed that the minimum sample size to detect non-normality in DDM is between 100 and 200 trials.

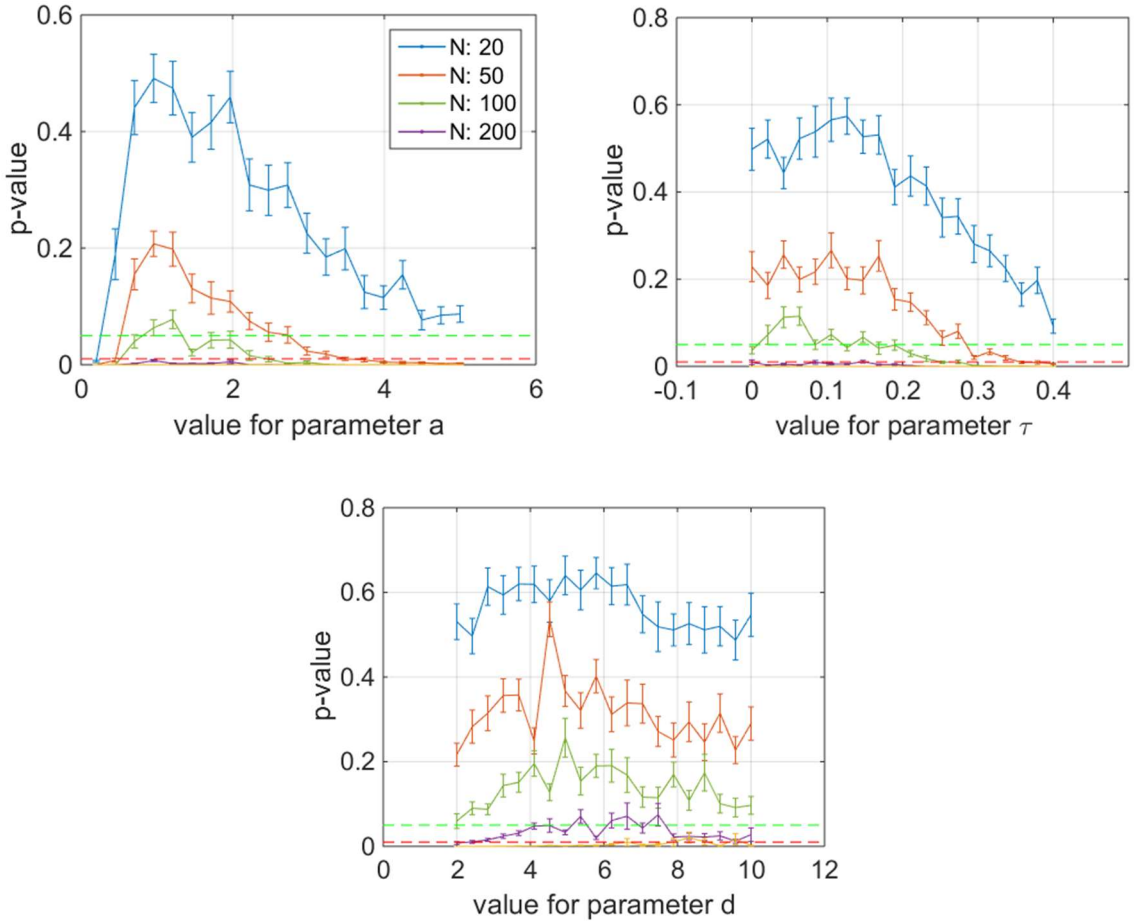


Figure 5.8

P-value for the Kolmogorov-Smirnov test with aggregated and standardised rate distribution to a normal distribution, with varying parameter a, τ or d. We used the Pure DDM and simulated 15 participants with 4 possible dataset size (N). When each one of these parameters were varied, the others were kept the same according to Table 5.4. We repeated each simulation 30 times. The vertical bar on each line indicates one standard error. Each line measure, informally, how likely is that the resulting distribution may be mis-identified as a Normal Distribution. The dashed horizontal green lines in each panel indicate the $p = 0.005$, the dashed horizontal red lines indicate $p = 0.001$. Note that, with small sample from each participants, most of the parameters will bring a mis-identification of the RT distribution (blue and orange line). Whereas both a and τ have a clear effect on the distributio nshape (increasing them will make the distribution less normal), the parameter d does not seem to have a relevant effect on the shape of the distribution, at least with the chosen set of parameters.

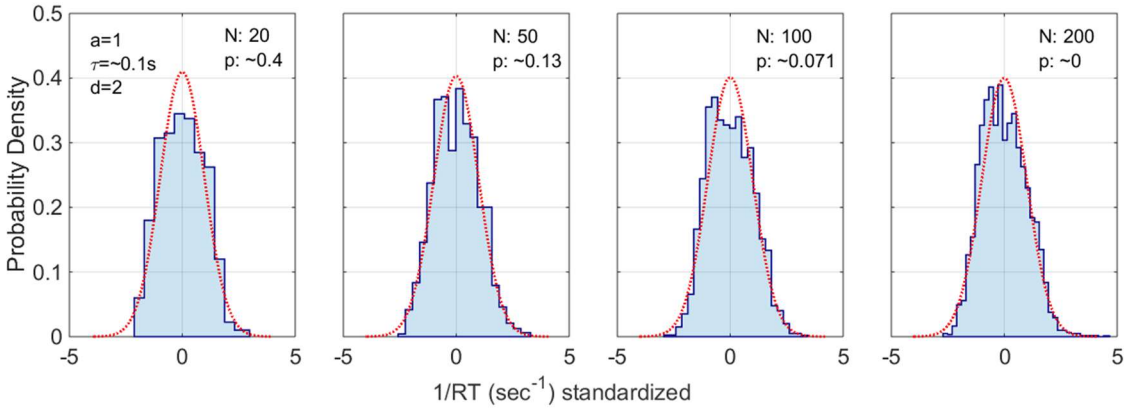


Figure 5.9

Typical resulting distribution from a set of simulated participants with parameters indicated in the left-most panel. Different panels indicate different sample sizes (N). Note that the sample size does not correspond to the size of the shown dataset, but to the size of the sample from each single participant (we used 15 simulated participants). The p-value indicated on each panel corresponds to the value obtained by a Kolmogorov-Smirnov test and corresponds to the probability that the sample is drawn from a Normal Distribution. Note how, by increasing the sample size, the distribution become clearly not normal and the test correctly detects it. However, with small sample size, like the one generally used in RT experiments, the test does not reject the null-hypothesis.

To compare this simulated results with the empirical observation, note that the number of samples for each condition was 60 for Experiment 4.1, and 40-90 for the Experiment 5.1 (remember that in this latter experiment the time was fixed so that the number of responses depended on participant's speed). According to our simulation, this sample size may not have been enough to differentiate between a real Normal Distribution and the distribution generated by a DDM. We, therefore, tried to collapse together all the standardised distributions from the Simple RT experiment in the previous Chapter, in the long mFP condition (the one that was less affected by fast responses), in order to have a unique dataset of 3600 observations and ran the Kolmogorov-Smirnov test on this dataset. This results in a $p = 0.0239$, meaning that the Normality can be rejected at $\alpha = 0.05$ but not at $\alpha = 0.01$. Again, this ambiguous results did not allow us to provide with a specific answer to the Rate-Hypothesis. The importance of the Rate-Hypothesis will be discussed again in the General Discussion in Chapter 9.

Chapter 6

Preliminary EXACT Paradigm

6.1 Introduction

In the previous two experimental Chapters, we have used a Classic RT paradigm to investigate the connection between experimental conditions (contrast, FP, RWI) and empirical behaviour, mostly in terms of RT distributions and accuracy. We have applied the two most successful sequential sampling models, the Pure and Extended DDM, and estimated the parameters of such models, in order to make inference about the decision strategy used by the participants. As explained in Chapter 2, we focused on 4 different decision strategies: Reward Rate (RR), modified Reward Rate (RR_m), Reward/Accuracy (RA) and Bayes Risk (BR). Three of these strategies are functions of the total non-decision time D_{tot} , and we modified this experimental parameter by either changing the FP or the RWI. Whereas we had obtained some interesting results by merely looking at the RT and accuracy, most of the information about what decision stage is affected by different conditions is given by the estimated parameters, which depend on the sequential sampling model used (DDM in our case). In this and the following Chapter, we develop a way to investigate decision strategies without having to rely on any particular assumption about the underlying mechanism. In particular, this Chapter will focus on the introduction of a new experimental paradigm, which will be exploited more deeply in the next Chapters.

6.1.1 Perceptual Process and Decision Rules

Decision-making can be conceptually broken down into two components: a perceptual process and a decision rule. The perceptual process is the mechanism that accumulates information and thus increases accuracy in order to decide between different alternatives, whereas the decision is used to decide when enough accumulation has been gathered (enough accuracy has been reached) for the

desired outcome. These two components are related: a decision rule depends on the accumulated accuracy, which depends on the perceptual model. For this reason, decision strategies and perceptual process have always been analysed together and the perceptual model used usually is the DDM (mostly the Pure version, Bogacz et al., 2006; Simen et al., 2009; Bogacz et al., 2010; Zacksenhous et al., 2010; Balci et al., 2011; for a comparison between DDM and linear ballistic accumulator in optimality see Goldfarb et al., 2014; for a review see Holmes and Cohen, 2014). The classic approach consists of (1) collect data from a classic RT experiment; (2) assume a perceptual model (DDM used almost universally); (3) based on the assumed perceptual process, test different decision rules (the most tested rules are RR and RR_m). We stress that the reason to assume a particular perceptual process is that it is necessary to have a way to relate the speed of response with the accuracy, that is, the speed-accuracy function $ACC(DT)$, where DT is the time of response in which an increase in accuracy is present (to be distinguished with the sensory-motor component, the part of the response in which no accuracy is accumulated). The $ACC(DT)$ is assumed to be dependent on experimental conditions such as stimulus strength and number of alternatives. Some qualitative predictions can be made without assuming any perceptual model. For example, by assuming a simple, monotonically increasing $ACC(DT)$, we can predict that the RT and accuracy will increase by increasing average trial length, as done in previous experiments. However, to obtain quantitative predictions, an exact equation of $ACC(DT)$ has to be found, and this is generally provided by the assumed perceptual process.

This leads to a problem when the researcher is interested in the decision strategy: what perceptual process should be assumed? As we have seen in **Section 2.2**, there is a myriad of perceptual processes, and they all generate different $ACC(DT)$ and, therefore, different predictions for different decision rule. When a decision rule appears not tocapture the empirical behaviour, the researcher is left to wonder whether this depends on the assumed perceptual process or on the decision strategy itself. Theoretically this problem could be solved by testing several perceptual models and check which one gives better results; however, given the high number of different models, and the fact that fitting them can sometime be a complicate process that involve choosing between several variants, this is rarely done in practice.

Moreover, it would always be possible to assume that an unknown perceptual process may give a better explanation for the data.

6.1.2 The EXACT Paradigm: Preliminary Version

For these reasons, we explored a new experimental paradigm in which there is no need to assume a perceptual model. The general idea is that the experimenter controls the speed-accuracy function ($ACC(DT)$) directly by "hard-coding" this function inside the experimental design. The participant is asked to press a key whenever she wants. A reward (in term of points) is given with probability $ACC(DT)$. We also introduce a punishment, given whenever the reward is not given. The participant's accuracy will therefore not depend on an endogenous process that accumulates information (like the DDM assumed previously), but on the $ACC(DT)$ function hard-coded within the paradigm: the task is an EXogenous ACcumulation Task (EXACT). The participant is informed that the accuracy function is always increasing, and thus waiting more increases the probability of obtaining a reward, but no other information about the $ACC(DT)$ is given (that is, the participant does not have direct access to the speed-accuracy function used in the task). We also limit the length of the condition. In this way, the participant has to decide between waiting more to increase the probability of obtaining a reward (corresponding to lower probability of getting a punishment), but resulting in less trials, or respond quickly, which would correspond in a lower probability of getting a reward (higher probability of a punishment), but would also result in a higher number of trials available. Finally, a pause between trials can be used. By using this design, all the parameters of the decision rules that we considered can be experimentally varied. Note that, with this paradigm, we can actually calculate the optimal response to earn as many points as possible by using RR_m and setting q (one of the parameter of the RR_m, referring to the emphasis on accuracy) equal to the punishment/reward ratio used (in terms of points): RR_m is the optimum strategy if the participants wants to earn as many points as possible. However, it is totally possible that participants subjective q is different from the experimental q. For example, participants may be very loss aversive, so they would have a strong emphasis on accuracy, or risk prone, and emphasize speed instead. In which case, if participants were responding according to RR_m, their

response will be suboptimal in terms of point, but will be optimal in terms of the subjective q that they were using.

Note that, whereas for the DDM there is a distinction between DT (the part of the response in which accuracy increases) and τ (part of the response without increase in accuracy), and $RT = DT + \tau$, in our case this distinction lose its meaning, as the accuracy increases from the beginning of the trial until response, even during the sensory-motor processing part. Thus, $DT = RT$, and we will refer to the speed-accuracy function as $ACC(RT)$ instead of $ACC(DT)$. Also, recall that the non-decision delay across trial, D_{tot}, is equal to $\tau + D$. In this paradigm the τ parameter does not have any meaning, and the delay across trial will be simply referred as D ($D_{tot} = D$).

Recall from **Section 2.7** that we distinguish 3 levels of analysis of optimal decision strategies, where we stress how quantitative prediction can be obtained only by assuming a particular perceptual process, but at the price of complex models and difficulty in the fitting process. By using this new paradigm, we can reach the same level of analysis, without having to fit any perceptual process, but by fitting the decision rules directly.

6.1.3 Exponential Speed-Accuracy Function

The main idea behind this paradigm is to hard-code the $ACC(RT)$ function within the experimental design, so that the researcher does not have to infer it by assuming a certain perceptual model. Similar to the experiment in the previous Chapter, this paradigm is designed as a point game: upon keypress at time t, the participant wins X reward points with probability $ACC(RT)$ and loses Y points with probability $1 - ACC(RT)$ (in practice, X points are assigned with probability $ACC(RT)$ and, if the points are not won, the tokens are taken away instead), where RT is the time from the beginning of the trial until participant's response. After a response, a pause may be inserted (to affect the parameter D of the decision rule), and then a new trial starts. To model $ACC(RT)$ we used a simple, monotonically increasing function, with enough flexibility to simulate the real $ACC(DT)$ for several experimental designs:

$$ACC(RT) = 1 + (\alpha - 1)\exp(-\lambda t) \qquad Eq.6.1$$

So that

$$RT(ACC) = -\ln\left(\frac{ACC - 1}{\alpha - 1}\right)/\lambda$$

Where λ controls how fast the function grows and α is the value at $t = 0$ (the probability of a correct response when no information is collected). Figure 6.1 shows several $ACC(RT)$ functions with different λ and α. These two parameters can be compared to different experimental designs in a classic RT task: different λ match to different task difficulty (by changing stimulus intensity, stimulus contrast, etc. where easier task corresponds to higher λ) whereas α may be related to the number of alternatives: $\alpha = 0.5$ represents a 2AFC, $\alpha = 0.25$ a 4AFC (this interpretation is not free of ambiguity, as will be explained in **Section 7.2**). Note that with this formulation it is also possible to explore more complicated design, such as $\alpha = 0.75$, which in a classic RT task may be designed as a task with 4 alternatives and 3 correct choices. For this chapter we set $\alpha = 0$.

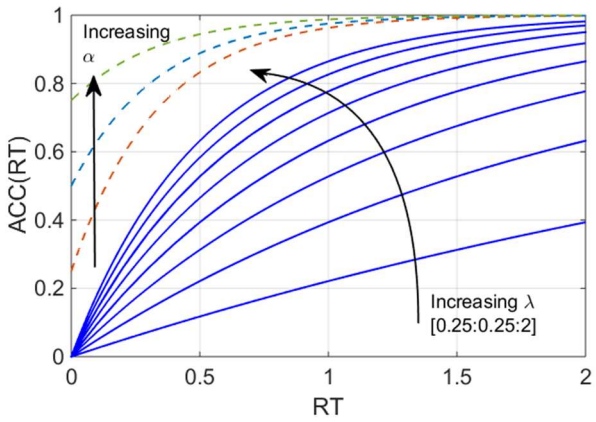


Figure 6.1

The speed-accuracy function ($ACC(R$) used for the preliminary EXACT Paradigm (Equation 6.1). Increasing λ corresponds to an increase in the steepness of the function, such that a certain level of accuracy is reached in less time. Increasing α corresponds to a shifting of the starting point, which, informally, corresponds to changing the number of alternatives in a Classic RT task.

6.1.4 Prediction for Exponential $ACC(RT)$

By knowing exactly the $ACC(RT)$ function, we can easily calculate the prediction for each one of the four decision rules used in this work. This was presented in **Section 3.6** and is reproduced here for convenience in Table 6.1 (note that, in this case, $DT = RT$ and $D_{tot} = D$).

Table 6.1 Decision Rules and Optimum Decision Time for exponential $ACC(RT)$

Decision Rules	Optimum decision time (RT^*), using $ACC(RT) = 1 + (\alpha - 1)\, exp(-\lambda RT)$
$BR = -\left[RT + q\left(1 - ACC(RT)\right)\right]$	$RT^*_{BR} = -\dfrac{ln\left(\dfrac{1}{\lambda q(1-\alpha)}\right)}{\lambda}$
$RR = ACC(RT)/(RT + D)$	$RT^*_{RR} = -D_{tot} - \dfrac{W_{-1}\left(\dfrac{exp(-\lambda D - 1)}{\alpha - 1}\right) + 1}{\lambda}$
$RA = RR - \dfrac{q}{D_{tot}}\left(1 - ACC(RT)\right)$	*No Explicit Form*
$RR_m = \dfrac{ACC(RT) - q\left(1 - ACC(RT)\right)}{RT + D}$	$RT^*_{RR_m} = -D - \dfrac{W_{-1}\left(\dfrac{exp(-\lambda D - 1)}{(\alpha - 1)(q + 1)}\right) + 1}{\lambda}$

Each one of these rules (left column) specified what is the "value" to maximise for each RT, given a certain accuracy function. The optimum response time RT^* is simply the time at which the value is maximized (right column of Table 6.1, see Appendix for derivation). Similar to what done **Section 2.8** where we analysed the predictions in terms of RT, accuracy and threshold parameter when the DDM is assumed, here we will analyse the prediction given the Exponential $ACC(RT)$ in Equation 6.1. The predicted optimum response (RT^*) values are shown in the left column of Figure 6.2, as a function of λ for different parameters values, and the right column shows the $ACC(RT^*)$ value (the reached accuracy level upon optimal responding). There is a monotonic relationship between D_{tot}, α, q and RT^*. The relationship with λ is more complicated: the optimum responses is 0 when λ is 0, as the stimulus does not give any information (the $ACC(RT)$ is flat). By increasing λ, the $ACC(RT)$ assume an increasing shape, so it become convenient to wait and accumulate information. After a certain point,

when the $ACC(RT)$ function is very steep, and the participant can start responding faster. With λ approaching infinity, the accuracy reaches 1 instantly, and the optimum responses goes to 0. This is the general pattern for RA, RR and BR. However, as Figure 6.2 shows, for RR_m the relationship is slightly more complex. For RR_m, in some particular conditions, RT^* goes to infinity when λ goes to infinity. This happens when $\alpha < \frac{q}{q+1}$ (see Appendix for proof). Coincidentally, the RR_m is particularly relevant as it is the only criteria that maximizes the gain for a task (included but not limited to the EXACT paradigm) with a fixed task duration, in terms of subjective utility value (Harris et al., 2014). This property allows RR_m to generate Piéron's and non-Piéron's shape (see next Section). We also notice the great similarities in terms of RT^* between RR_m and RA. However, note how RA produce a very peculiar $ACC(RT)$ shape.

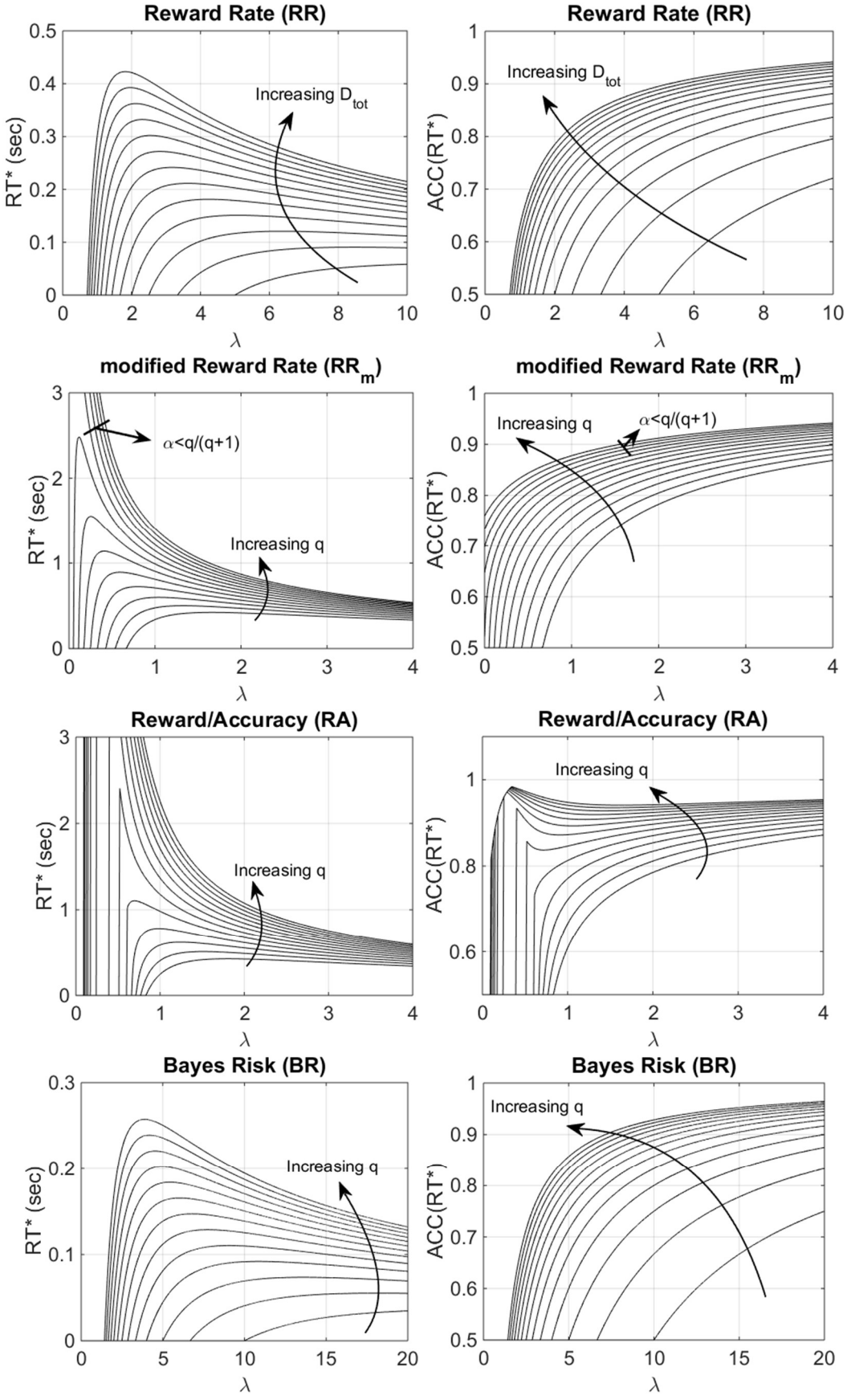

Reward Rate (RR)
RT* (sec)
Increasing D_{tot}
λ
Reward Rate (RR)
ACC(RT*)
Increasing D_{tot}
λ
modified Reward Rate (RR_m)
RT* (sec)
α<q/(q+1)
Increasing q
λ
modified Reward Rate (RR_m)
ACC(RT*)
Increasing q
α<q/(q+1)
λ
Reward/Accuracy (RA)
RT* (sec)
Increasing q
λ
Reward/Accuracy (RA)
ACC(RT*)
Increasing q
λ
Bayes Risk (BR)
RT* (sec)
Increasing q
λ
Bayes Risk (BR)
ACC(RT*)
Increasing q
λ

Figure 6.2

Optimal response time, RT^* **(left column)** and $ACC(RT^*)$ **(right column)** vs. λ for the four decision rules, with different parameters q and D_{tot} . $ACC(RT^*)$ indicates the optimum level of accuracy that has to be reached upon responding to maximize a particular decision rule. The arrows indicate how the shape changes when the D_{tot} or q parameter is increased. The q parameter represents the subjective punishment/reward ratio; the D_{tot} parameter represents the total non-decision time of the trial (sum of experimental delay and sensory-motor component, $(D + \tau)$. Increasing q or increasing D_{tot} corresponds to an increase in the optimum response time (RT^*) and a consequent increase of $ACC(RT^*)$. For the modified Reward Rate (second row from the top), two very different RT^* shapes are produced with different combinations of q and α (the starting point of the gauge). We refer to the shapes indicate by the black line as Piéron's shape, the others as non-Piéron's shape.

6.1.5 Piéron's and non-Piéron's Shape with Exponential $ACC(RT)$

We showed in **Section 2.10** that RR_m, with some peculiar parameters, produces a relationship that is practically indistinguishable with Piéron's Law. Recall that Piéron's Law (**Section 2.9**) relates the stimulus strength with the mean RT ($mRT = \gamma + kI^{-\beta}$). Here, the stimulus strength can be interpreted as the λ parameter in Equation 6.1 and thus $mRT = \gamma + k\lambda^{-\beta}$). However, for some RR_m parameters, and for all the other decision rules, the relationship had a Piéron's shape only if considered from a point further than the maximum of the DT^* curve (Figure 2.5 and 2.6 in **Section 2.10**). Now that we have presented the mathematical formulation of the $ACC(RT)$, we can more precisely specify what parameters need to be changed to produce Piéron's and non-Piéron's Law.

In the Appendix, we show that when $\alpha < \frac{q}{q+1}$, RR_m produces Piéron's-shape curves, whereas when $\alpha > \frac{q}{q+1}$, the relationship does not follow a Piéron's Law for all level of stimulus strength λ: with very low level of λ the optimum response RT^* decreases, and Piéron's Law does not hold. For all these cases, a Piéron's Law can still be fitted, if low values of λ are excluded (Figure 2.5). This is particularly interesting, since in all the past experiments with Piéron's Law focused on stimulus strength (see Section 2.9 for the literature on Piéron's Law), the non-Piéron's shape was never observed, even though is predicted with almost all experimental conditions and decision rules. Our framework can easily explain why this has never been reported, at least for RR_m: recall that α may be interpreted as the number of alternatives (so that a 2AFC task corresponds to an $\alpha = 0.5$), and q corresponds (for RR_m) to the subjective punishment/reward ratio. This means that, in a 2AFC task, the non-Piéron's shape will be obtained when $q < 1$, which means that the punishment would need to have a *lower* subjective

magnitude than the reward. This would require the researcher to give an extremely high reward, or decrease the penalty for incorrect responses. In all the Piéron's studies, however, incorrect responses are highly penalized, and we believe that this is the reason why the non-Piéron's shape has never been observed. In Chapter 7 we will see how, within the EXACT Paradigm, both shape can be generated by manipulating α. In Chapter 8 we investigate the effect of manipulating q with the EXACT Paradigm and compare it with a similar Classic RT paradigm to see if, as predicted by RR_m, we can produce Piéron's and non-Piéron's shape by varying q.

As a final remark, we stress again the fact that, though theoretically only RR_m can produce a Piéron's shape, observing such shape does not indicate uniquely that the participants are using RR_m: RR, RA and BR may also be fitted to Piéron's Law, as until a certain point they may appear to increase with decreasing stimulus strength. It may be difficult to distinguish between a shape that is really growing to infinity when λ goes to 0, and a shape that keeps growing with extremely low values of λ, before decreasing suddenly. Observing a Piéron's shape cannot be the definitive indicator that the RR_m rule is employed, and a quantitative fit to the model is still necessary.

6.1.6 Comparison with Classic RT Task

The EXACT Paradigm can be interpreted as an abstract version of a Classic RT task. The use of the exponential, monotonically increasing $ACC(RT)$, makes it possible to perform several comparisons between this and Classic RT tasks, and in fact several predictions are similar to the ones generated by a DDM. However, there are some relevant differences: firstly, the EXACT Paradigm does not allow a real choice between different alternatives. This means that investigating biases towards alternative is not possible (see **Section 7.2**). Secondly, according to most perceptual models, the process of increasing information is noisy, whereas in our case is a continuous increasing function, which make us lose some important phenomena, such as investigating errors that are slower than correct responses (e.g. Luce, 1986, Ratcliff & McKoon, 2009, Leite, 2012): in our paradigm, on average, correct responses are always slower. In spite of these limitations, we still believe that this approach may lead to a better analysis of decision rules. Plus, these limitations are balanced by the possibility of having designs that are

impossible with Classic RT tasks: e.g. the already mentioned $\alpha > 0.5$ and the possibility of varying $ACC(RT)$. Furthermore, note how by using a standardised $ACC(RT)$ we do not have to be worried about the noise introduced by different participants' visual skills: even assuming the same $ACC(RT)$ function across different individuals, in a Classic RT brightness detection task, the same stimulus may produce different steepness (in terms of their accuracy functions) depending on different participants' visual acuity (we are going to see this effect in the second experiment of Chapter 8). By using a standardised $ACC(RT)$, we overcome the problem of individual differences in task skills/visual ability, and we can focus on the analysis of individual decision strategies. Finally, we want to stress how the EXACT Paradigm, despite being similar and comparable with Classic RT task, is not supposed to mimic exactly such tasks. The main aim of this paradigm is to provide an experimental framework to investigate decision rules. Thus, we do not deem the fact that some phenomena observed in Classic RT are not reproducible as an invalidating flaw for the EXACT Paradigm.

EXPERIMENT 6.1

Compared to the previous experiments, the study in this Chapter has to be seen like an "exploratory" investigation. Since we were using a new experimental design, our main focus was not only to test some particular hypothesis, but to see in general the feasibility of such design, how participants responded to it, and if it can be consistently applied to investigate complex decisions. We kept the design very simple by having two conditions in which we vary uniquely the delay. Similarly to the past Chapters, we are going to check whether participants change their response based on the average trial length (D), and we do this by varying the time across trials. Also, as we used only two conditions, we refrain from fitting specific decision rules. More complex design, coupled with quantitative fitting, will be presented in the next Chapters. We are particularly interested in stability of responses, and if participants are generally responsive to different experimental conditions. We will discuss the limitations of this design and suggest an improvement upon it, which will be implemented in Chapter 7 and 8.

6.2 Methods

Instructions. To better understand what type of instructions to use, we firstly ran a pilot version of this study in which the participant were told that they could press the key at any time, and a reward was given to them with a certain *increasing* probability and, when a reward was not given, a punishment was given instead. The reward probability corresponded to the $ACC(RT)$ presented in Equation 6.1, but they were not informed about the exact shape of this function. The problem with this approach was that participants often thought that the reward and punishment followed some complicate schedule that they had to learn (for example, they needed to response one time fast, the second time slow, alternatively, along all trials). In order to solve this problem, we tried to convey a different mental framing, by using different instructions: we told participants that an invisible stimulus appeared after some time, and that they had to respond whenever they thought the stimulus appear[*]. This design conveyed the idea that the more they waited, the more likely they were to obtain a correct response, without the presence of any hidden pattern. The idea was that, after several trials, and after obtaining several reward and punishment, they could learn the function that ruled the task.

We recruited 30 (12 males, 18 females) participants to take part in this study. Each participant underwent 2 training sessions and one main experimental session.

Training Sessions. In order to reinforce the idea that participant had to respond to an invisible stimulus, and not trying to figure out a random pattern concealed to them, we ran the first training session where a real stimulus was present. In the first training session, lasting one minute, participants

[*] Note that giving a reward at time RT according to a probability specified by $ACC(RT)$, or giving a reward for responses made after time X randomly distributed according to $\frac{dACC(t)}{dt}$, are exactly equivalent, as $ACC(RT)$ can be seen as a cumulative distribution (in our, a cumulative exponential distribution). Thus, the design that we explained to them was mathematically equivalent to the one that we were using in the code.

were asked to respond to a stimulus (a circle appearing on the monitor) as fast as possible. The circle appeared randomly according to an exponential distribution with mean equal to 2.5seconds ($\lambda = \frac{1}{2.5} = 0.4,\ \alpha = 0$). After the participant responded, the next trial began instantly ($D = 0$). Overall, this was just a Simple RT task with a stimulus well above threshold. In this way, we tried to convince participants that the appearing of the stimulus was random and did not follow any complex pattern, as they thought in the pilot study. After this, in the second training session, we told participants that now the stimulus was not going to be visible (they had to estimate when it had appeared), and press a button when they think it had so. In this case we, set $\lambda = 1, \alpha = 0$. Upon participant response, the reward was given (points assigned) with probability equal to $ACC(RT)$, where RT was the time from the beginning of the trial until the response. When the reward was not given, the punishment was given instead (points taken away). For this training session we used a delay D between trials of 1 second. This training session lasted one minute.

Procedures and Materials. The main study was similar to the second training session: upon response, participants received a reward with probability $ACC(RT)$, otherwise they received a punishment. As before, $\lambda = 1, \alpha = 0$. We used two delay conditions: $D = 0s$ or $D = 2.5s$, and ran one 12 minutes session for each condition. The reward and the punishment were set to 1 point, and each participant started each block with 25 points. With these parameters, if participants were trying to maximize the amount of point earned during each session, then the difference in RT between the two D conditions should be of 0.8 seconds (calculated using the RR_m function). We used a timeout of 10 seconds, whereupon a punishment was given, but it was rarely reached by any participants. To increase incentive, the total amount of points earned was exchanged for money at the end of the experiment, at the rate of 2 points for 1 penny. If any participant went below 0 points, their score was set to 0. We also used a negative sound feedback when an error was made. The point earned or lost in the last trial was showed at all time on the screen (e.g. +1 or -1), along with the total score for each session. These two numbers were shown in red if the last trial resulted in a punishment and in green if it resulted in a reward. The monitor also showed a progressing gauge aimed to help the participant to keep track of the passing

of time. The gauge was empty at the beginning of each trial, and filled up linearly with time, reaching its maximum at 10 seconds. The gauge did not have any relationship with the $ACC(RT)$ (as in the real EXACT Paradigm, presented in the next Chapter). Finally, the screen also showed the time left in the session (in minutes). In the condition with $D > 0$, from a trial to another we introduced a delay of D seconds. During this delay, the gauge disappeared and a "Please Wait" writing appeared instead. All the other information stayed on the screen. An example of the monitor screen seen by the participant during a trial is shown in Figure 6.3. In this case, we can see that the trial started from some time (as the gauge is not completely empty), with 2 minutes left in the session, 23 points in total for this session and the last trial resulted in an error with a punishment of 1.

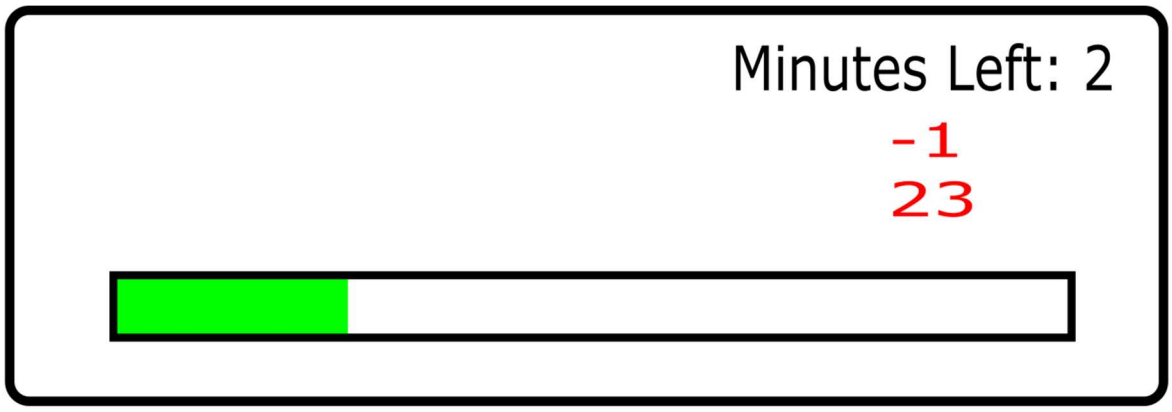


Figure 6.3

Illustration of what is shown to the participant on the monitor screen. The gauge linearly (the green horizontal bar) increases with time until the timeout is reached. The red numbers indicate the outcome of the previous trial (on top) and the total score for the current condition (on the bottom). The number of minutes left in the condition is also indicated.

Analysis. We calculated the optimum response for each condition, in which optimum is defined as earning as many points as possible. The optimum response RT^* is obtained by calculating the maximum of RR_m with $q = 1$ (recall that, in RR_m, q corresponds to the punishment/reward ratio, which in this experiments are both equal to 1). By applying similar reasoning, we can calculate the expected score given the optimum performance. As this is a probabilistic task, participants may sometime earn slightly more than the expected score. Given an optimum response RT^*, we can obtain the optimum accuracy $ACC(RT^*)$ by applying Equation 6.1.

We will analyse the response both in terms of RT and in terms of $ACC(R$ (the accuracy corresponding to participants' response). For the RT, we calculated the median of the median of responses for each participant and each condition (mdRT). For the $ACC(R$, we calculated the median of the median of the accuracy for each response, participant and condition.

6.3 Results

We excluded one participant as a clearly uncooperative (his responses were extremely long, over 3 standard deviations over the mean). For the remaining participants, we compared the optimum expected score (expected amount of point at the end of each condition, given that participants were actually trying to maximise the amount of point earned, see Methods above) for each one of the two conditions with the effective gross by each participant (Figure 6.4). Note that in the figure we add a horizontal noise to make easier to distinguish amongst different participant's data points. Most of the participants appeared to be far away from the theoretical optimum for the first D condition, but closer for the second condition.

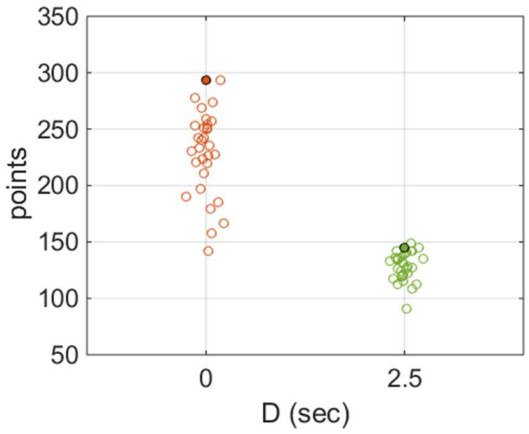


Figure 6.4

The empty circles represent the points obtained for each participant for the two D conditions. The filled circles represent the expected average score with optimum performance (calculated with RR_m). Noise on the horizontal axes is added to increase readability. Participants appeared to be responsive to the changing of the condition, but they were consistently sub-optimal.

6.3.1 Response to Feedback

Another way to analyse participant's response to the task is to observe the change in their response after a punishment or after a reward. As the participants knew that the more they waited, the more likely they were to receive a reward, we expected a slowing down of RT after a punishment and a speed up

after a reward. For each participant and each condition, we calculated the difference between RT corresponding to the response after a reward and the previous RT, and compute the change to when a reward or punishment was received. The results show that, as predicted, there was an average speed up of responses of about 0.025seconds after a reward. By applying the same procedure to the responses after punishments, we obtained a slowing down of about 0.45 seconds. We did not found any relevant difference between the two D conditions. The difference in absolute terms of the speed up compared to the slowing down can be due to the much higher number of obtained rewards than punishment: this made possible for the participants to adjust more finely the response after a reward, and more decisively after a (rare) punishment.

It is possible that the change in response after a reward/punishment depended on the RT length itself. For example, a participant that is producing a series of very long RT, after a punishment, may not adjust her response as much as a participant which produced a series of very short RT. We ran this analysis by dividing the RT into 10 bins for each participant and calculated the average adjustment in response after a reward/punishment for each bin. The small number of punishments for some participants and some bin range made this analysis problematic. Furthermore, as different participants have different RT ranges (and the bin were calculated respectively to each participant's range of response), the analysis for the reward was also difficult to interpret. By visual inspection of individual participant's change of response across different time bin, we could not detect any interesting pattern. It appears that either a more sophisticated approach is needed here, or participants were not actually adjusting their response based on the speed of the response itself.

6.3.2 Response Signal Analysis

As this was an exploratory study, we were interested in checking the stability of responses along each one of the two conditions. Firstly, we analysed the RT and $ACC(RT)$ for each participant and each condition along the time of the trial (we call this time series the response signal). Not only the results were remarkably variable from a participant to another, but they were variable also within participants. For example, the top panels of Figure 6.5 shows a fairly stable response signal (in terms of RT

and $ACC(RT)$) for $D = 0s$, with a small non-linear trend when $D = 2.5s$. However, bottom panels of Figure 6.5 show a very different pattern, with participant suddenly shifting response speed for $D = 0s$, and following a more complex strategy for $D = 2.5s$. There are several different trends and complex patterns in the dataset obtained in this study. We applied the sequential t-test analysis of regime shift by using the STARS method (Rodionov and Overlan, 2005, see Chapter 3 for Methods) which returns the regime shift index for each participant and each condition, indicating the number of shifts in the time series and the magnitude of those shifts (scaled by the standard deviation). This gives us an idea of how often participants changed their strategy and how stable their response was. We can then compare these values with those from the Classic RT experiment and from the other EXACT paradigm studies (in the next Chapters). Note that we applied STARS on the RT, but applying to the $ACC(RT)$ led to very similar results. Figure 6.5 shows the position in time and the magnitude of the detected regime (magenta lines). By visual inspection it appears that STARS correctly classified the response time for both signals, detecting less than 5 shift changes for each session in the top panels of Figure 6.5a and many more for the signal in the bottom panels of Figure 6.5.

The result of this analysis revealed that there were many more changes of strategy for the $D = 0s$ condition (average num. of detected shift index across individuals was 11.23, STD=4.454) than for the $D = 2.5s$ condition (with an average of 5.66 shift index, STD=2.630). The magnitude of the shift (calculated by averaging the absolute value of the SI grouped across all participants) was more variable, with a mean of 0.72 and 0.82 and a STD of 0.69 and 0.81 (for $D = 0s$ and 2.5s, respectively).

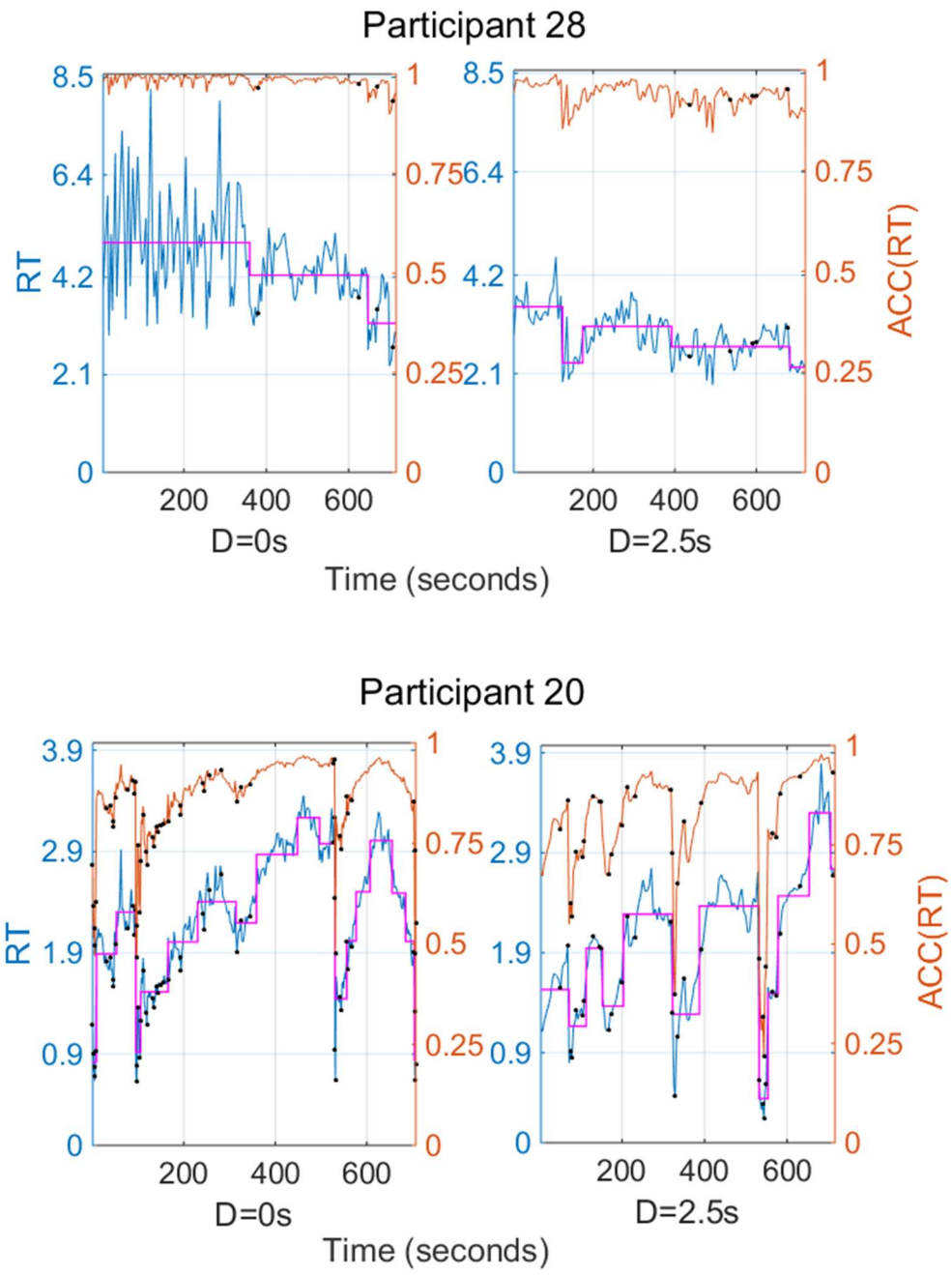


Figure 6.5

Response Signal for two selected participants, for the two conditions. The blue lines represent the response in terms of RT (left vertical axes), the orange lines represent the response in terms of $ACC(RT)$ (right vertical axes, see Equation 6.1 for the relationship between RT and $ACC(RT)$). The blue dots on the lines correspond to error responses. The magenta lines are the result of STARS method, and they represent the detected shifting in strategy. The top panels refer to a participant with relatively stable response signal, whereas the bottom panel show a participant with trends, and frequent and sudden shift of strategies.

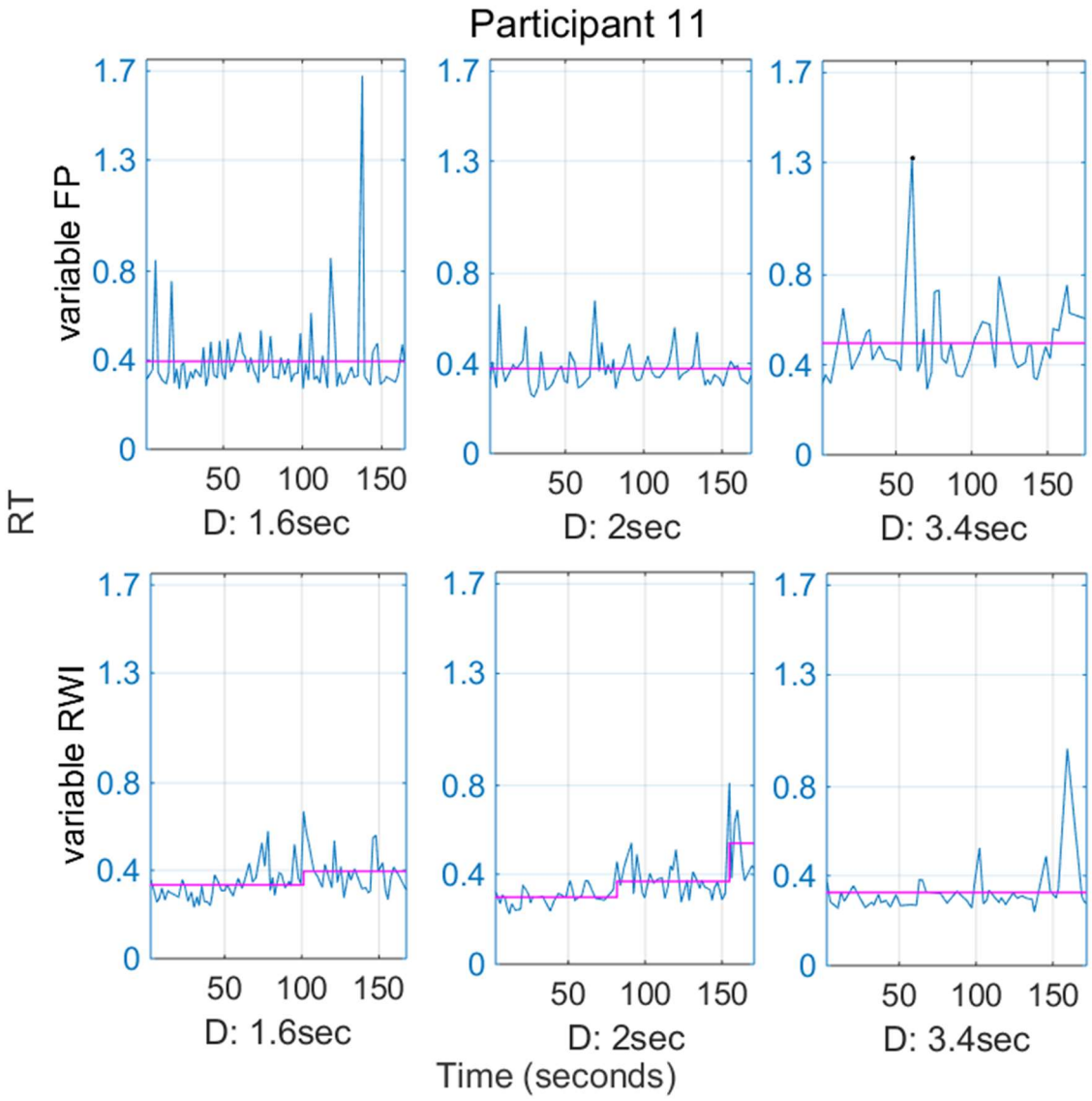


Figure 6.6

Response Signal (in terms of RT) for a typical participant in Experiment 5.1 (blue lines). The magenta lines correspond to the outcome of the STARS method, indicating the detected shifts in the response. Comparing this figure with Figure 6.5, it is clear that in the experiment in the present Chapter there is a higher amount of response shifting.

Note that these response signals are very different from what we obtained with the Classic RT tasks in the previous Chapter, as those were almost uniquely stable, without any particular trend or rapid shift.

This was again confirmed by the STARS method applied to the signal for the experiments in Chapter 4 and 5, which detected on average less than on shift for each participant and condition, with a magnitude of the shift lower than 0.5 (Figure 6.6 shows a typical participant from Experiment 5.1).

Note that the STARS method revealed information about number and magnitude of shifts in the response signal, but does not tell us anything about the presence of trends/pattern in the time series. By visual inspection of the data, we were not able to notice any general pattern, as participants' strategies looked too different from one another to draw any sensible conclusion. Therefore, we tried to collapse the data together. We divided the condition length (in seconds) in 20 bins (from 0 to 720 seconds) and calculated, for each participant, the average response within that bin. We then averaged the responses across all participants. In Figure 6.7 we show the result of this analysis for both conditions, where the shaded area represents one standard deviation. We fitted a simple linear model to the data and calculated the correlation values. There is indeed a clear decreasing relationship between time on the trial and RT, but only for $D = 0s$ condition ($p < 0.005$). This may indicate that, at the beginning of the trial, participants started with some prior expectation about the trial parameters, and then slowly adjusted their response in light of the real experimental parameters. In our case, for example, participants may have expected a delay across trial higher than $D = 0s$, and then they had to adjust their response to correct for it. Note that, by running the same analysis on the Classic RT tasks in Chapter 4 and 5, we did not find any trend.

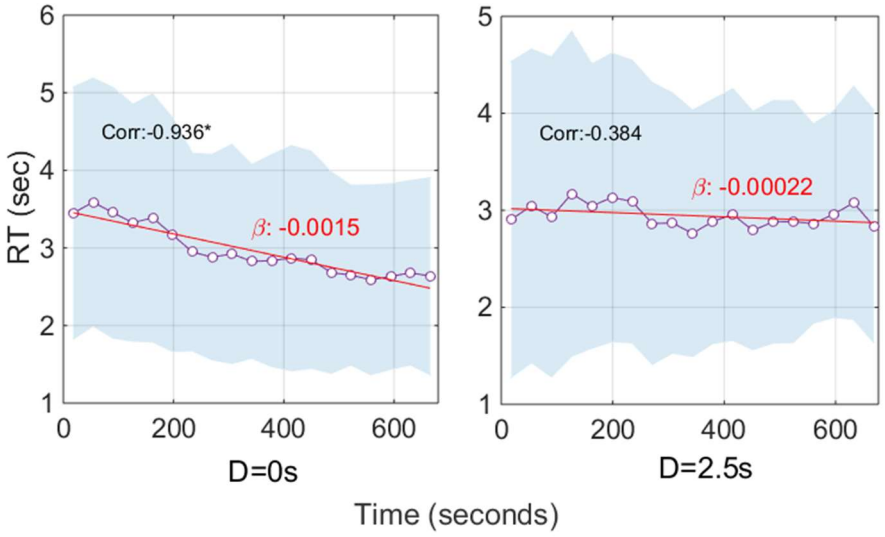


Figure 6.7

General trend for the Response Signals (in term of RT), obtained by averaging the binned responses in time (see text for details). The red lines correspond to a linear fit with slope indicated in the figure (β). There is a significant correlation between

time and response time only for the first D condition, as indicated by the asterisk in the left panel. The shaded area represents one standard deviation.

We also ran a similar analysis on the standard deviation, calculating the mean RT for each bin, we computed the standard deviation. This analysis can reveal whether participants were decreasing their variability along length condition. For both conditions we found a negative, weak correlation between time in the condition and standard deviation of the response, meaning either that participants were not adjusting their response along the duration of the experiment, or that the condition did not last enough for getting their response on a steady state.

6.3.4 Effect of Delay Across Trials

The results in terms of median of median RT (mdRT) and accuracy of response ($ACC(RT)$) are shown in Figure 6.8 (blue lines). The median RT and $ACC(RT)$ distribution for each participant is shown vertically. Participants appeared to increase their RT with increasing D but the difference, calculated on the median RT was not statistically significant ($F_{1,29} = 2.108, p = 0.157$). Similarly, the effect on accuracy of response was not significant ($F_{1,29} = 1.936, p = 0.175$). The optimum response (calculate based on RR_m with $q = 1$) is shown for comparison (red crosses). The theoretical difference in terms of RT between the two D conditions is 0.8 seconds, whereas we observed a much lower difference of only 0.18seconds. The optimum score is close to the empirical one for the $D = 2.5s$ condition, but not for $D = 0$. For $ACC(RT)$, the theoretical difference is 0.1, the empirical one is ~0.02. Recall that we have observed, in analysing the collapsed trends, that participants were slowly adjusting (decreasing) their response for $D = 0$. This could be a possible reason why no significant difference was observed between the two conditions. We took that into account by eliminating the first N trials from the dataset to see how the mdRT and median $ACC(RT)$ changed after the initial period. We varies N from 20 to 100 in steps of 20, and the results are plotted as black lines in Figure 6.8. As expected from the trend analysis, the mdRT and median ACC(RT) for the $RSI = 0s$ slightly decreases, whereas for $D = 2.5s$ it remains more stable. However, by running an ANOVA on these modified dataset, the median RT were still not significantly different for the two conditions.

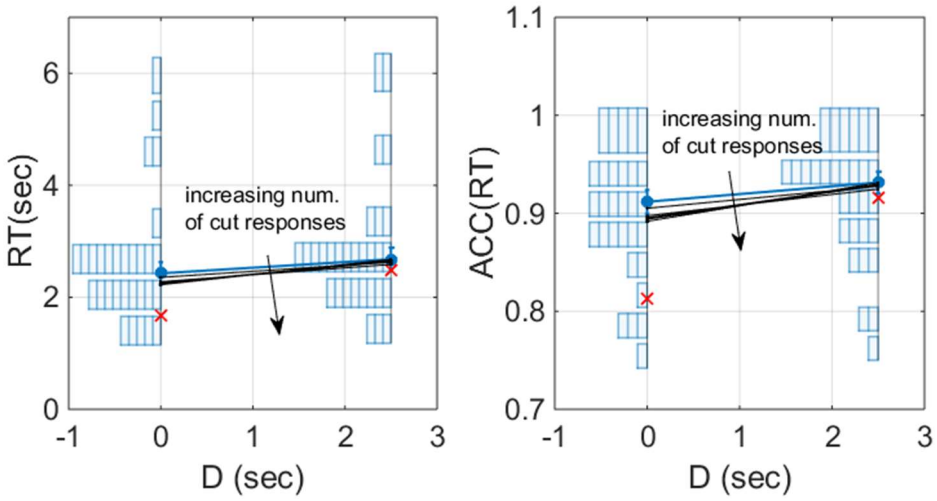


Figure 6.8

Median of median RT (mdRT) **(left panel)** and median of median $ACC(RT)$ **(right panel)** for the two D conditions (blue circles, the blue lines connect the two conditions). We overimposed vertically the frequency distribution of the median RT for each participant (blue blocks). The red circles represent the optimum response (calculated with RR_m). Even though there is a change for both RT and $ACC(RT)$, the increase is not significant. In particular, the response for the $D = 0s$ is significantly higher than the optimum response. We iteratively cut initial responses to take into account any possible learning period, but this did not seem to make any relevant difference (black lines).

We grouped the 30 participants in three classes according to their performance (poor, average and good Performance, corresponding respectively to the worst third, the middle third, and the top third in terms of point earned). Similarly to the collapsed results, for each one of the performance group mdRT appeared to increase with increasing D, but again the difference was not significant for any of the three groups (Figure 6.8., for the poor, average and good performance respectively $F_{1,9} = 0.082, p = 0.781$, $F_{1,9} = 0.523, p = 0.488$, $F_{1,9} = 3.335, p = 0.101$). Unsurprisingly, the group that performed best (earned the most point) also had faster RT in all conditions (Figure 6.9).

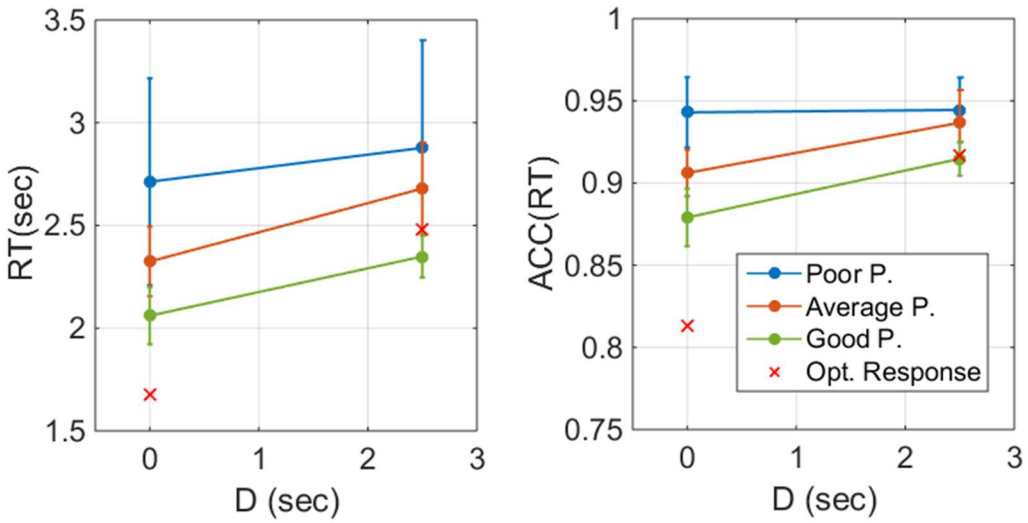


Figure 6.9

Median of median RT for each performance group. Similar to the grouped responses in the previous Figure, the D condition appear to increase the responses, but the effect was not significant. As expected, better performance groups corresponded to faster responses. The red crosses indicate the optimum responses calculate with RR_m.

We were also interested in analysing the distributions to check similarities with Classic RT paradigm. Similar to distribution in the Classic RT tasks, the response distribution with $D = 0s$ was slightly skewed on the right (however, the skewness value is much lower in this case than in Classic RT paradigm: compare it with distributions in Figure 5.4). Despite the non-significance of mdRT between the two D conditions, the distributions for the $D = 2.5s$ condition and the distribution for the $D = 0s$ condition appeared clearly different: the latter did not present almost any skewness (Figure 6.10). Difference in time distributions will be discussed in **Section 7.10**. We analysed the distribution in the Rate Domain in order to compare it with the distributions obtained with the Classic RT design. Although the distributions in time domain were not extremely different with what normally observed (at least for the $D = 0s$ condition), the distributions in the rate domain were remarkably different to what normally observed in Classic RT tasks, and they were very dissimilar to a Normal Distribution (see Figure 6.11). Also, applying the STD-QRI method did not affected much the distributions (this is a further proof that this method does not generate a Normal Distribution if the original distribution is not normal).

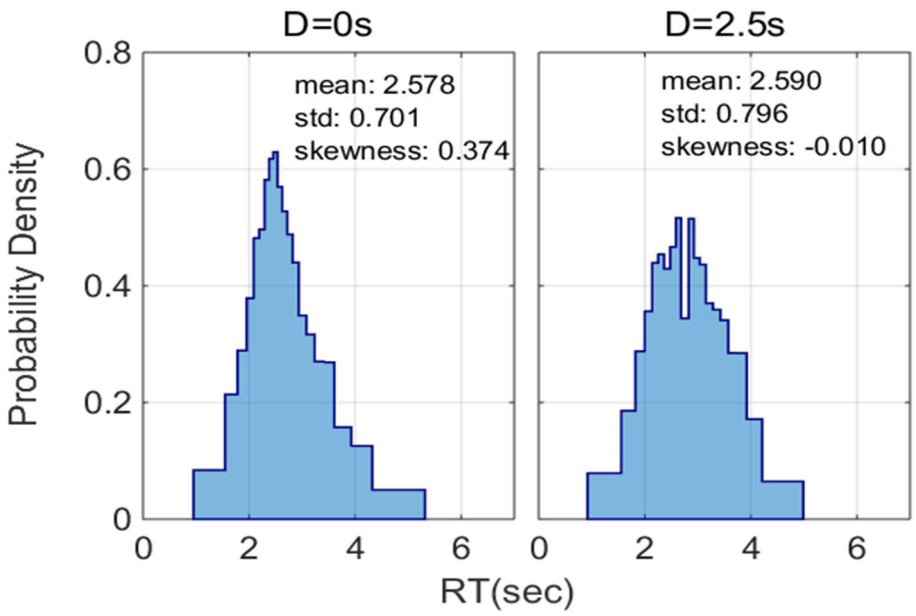


Figure 6.10

Vincentized RT distributions for the two D conditions. As in the distributions for Classic RT tasks, the distributions for this paradigm appeared slightly skewed on the right side, but only for the $D = 0s$ condition.

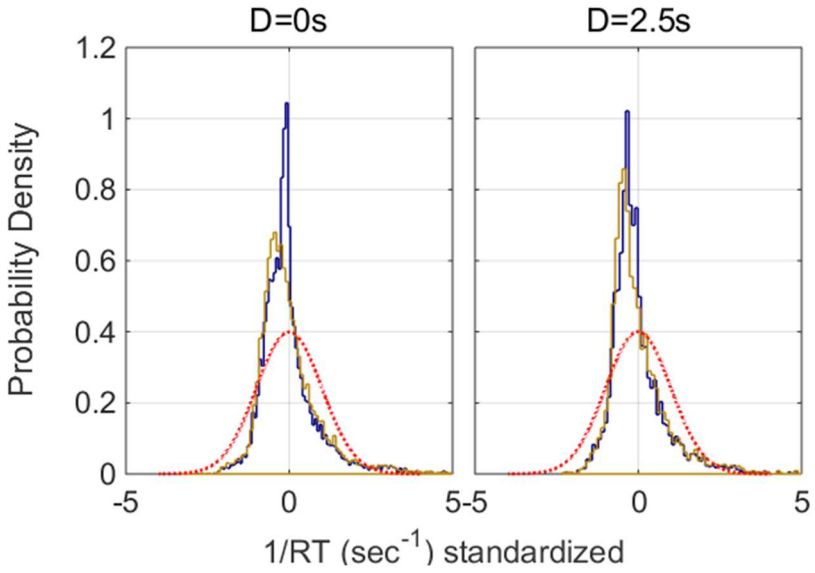


Figure 6.11

Aggregated and standardised Rate Distributions, before and after applying the STD-IQR method for eliminating contaminant data (blue and yellow line respectively). The dashed red line indicates a Standard Normal Distribution. Conversely to what obtained in the previous Classic RT experiment, here the responses are clearly not-Normal Distributed.

6.4 Discussion

The result of this preliminary experiment taught us several important aspects of this new paradigm. Firstly, responses are much more variable in this paradigm than in Classic RT tasks, with sudden shifting and complex patterns. However, this may be because participants had to estimate the correct $ACC(RT)$ shape, which may be an extremely difficult task. In the next Chapter, with the complete version of the EXACT Paradigm, we are going to explicitly show the $ACC(RT)$ on the screen so that participants do not have to estimate it. In doing so, we will obtain more stable response (however, as we will see, these will still not be as stable as the responses from a Classic RT task).

Also, participants reported that they found the negative sound feedback extremely annoying. It is possible that this contributed in making their response more loss-aversive, increasing the size of adjustment for punishment, and decreasing the difference across the two D conditions.

We did not notice any significant effect of D on responses, confirming previous studies' results. The pattern was however in the expected direction, with responses that increased with increasing delay across trial. This was also true for each individual performance group. Is it possible that participants were adjusting their optimization response based on D , but in a much lower degree than expected, so that the difference did not resulted to be significant. However, as found in the previous Chapter, the difference in responses across difference D conditions may be due to difference participant's temporal preparation, which may be lowered by higher D conditions. This implies that participants are not changing their response based on an optimization process, or that the optimization process is not function of the delay across trial. We reached a similar conclusion based on experiments in previous Chapters.

Another relevant difference regards the rate distributions, which looked extremely different than those from a Classic RT task. As Chapter 7 and 8 will confirm, this is a stable feature of the EXACT Paradigm: whereas RT distributions look reasonably similar to Classic RT tasks (even more with next

studies), rate distributions are very dissimilar from what we observed in the previous experiments, that is: they are not approximately normal, in any of the several experimental conditions. The reason behind this may lie in the variability of the speed-accuracy shape across trials. In this paradigm, the $ACC(RT)$ is supposed to be the same across all trials for a certain condition, regardless of participant's real accuracy. However, in Classic RT tasks, if we take attention lapses, or noise in the sensory system into account, we see how $ACC(RT)$ may slightly vary from a trial to another. A fluctuating $ACC(RT)$, coupled with some decision strategy, may be the reason of normal distribution in the rate domain (see Harris et al., 2014).

Note that in this exploratory experiment, we deemed unnecessary to proceed with a complete fitting routine for the four decision strategies: by having only two conditions, most of the decision rules produced a perfect fit, with very small differences from one another, so that comparing the model seemed meaningless. In the next Chapter, when employing several conditions within the context of the complete EXACT Paradigm, we will be able to test quantitative prediction and compare more precisely the goodness of fit of each one of the four models, in order to exploit the power of the paradigm itself: being able to fit precise quantitative prediction without having to assume any perceptual process.

Chapter 7

The EXACT Paradigm

7.1 Introduction

In the previous Chapter, we introduced a new experimental design in which the increase in accuracy did not depend on a participant sensory perception (as usually happens in Classic RT task) but is hard-coded into the experimental design, and perfectly known by the researcher. The main feature of this paradigm was that the speed-accuracy function ($ACC(RT)$) was kept hidden to the participant. The results from this study revealed that the responses were much noisier than Classic RT task, and they may present trends, conversely to responses coming from Classic RT tasks. We believe that such noisy responses and trends depended on the fact that participants were continuously trying to estimate the $ACC(RT)$ function and, at the same time, trying to adjust the response in order to find the optimal response. This resulted in frequent, sudden strategy shifting for most of the participants, or slow but steady trends for some others. Hiding the $ACC(RT)$ function not only makes responses more complex, but it also seems an unnatural feature. In fact, it is commonly believed that in Classic RT experiments participants base their response on the perceived accuracy, which they may have at least partially access to (Bogacz et al., 2006). Therefore, in the EXACT Paradigm that we present in this chapter we made sure that participants also had access to the $ACC(RT)$, by showing a graphic representation of it on the monitor in the form of a progressing gauge. As in the previous Chapter, the name comes from the fact that the accumulation of information is exogenous to the participant (EXogenous ACcumulation Task). As before, we show on the screen only the minimum information required to perform the task: the points earned/lost in the previous trial (setting the colour equal to red if corresponded to a punishment, or green for a reward), the total score for the session, the time left in the condition in minutes, and a progressing gauge. Note that in the previous version we also had a progressing gauge which however was used to help participant to track the time spent in each single trial (Figure 6.3). In this case, the gauge is used

to represent the $ACC(RT)$ directly: it is empty when $ACC(RT) = 0$**,** is full at $ACC(RT) = 1$**,** it starts at a point equal to α (for example, with $\alpha = 0.5$, it starts half full) and it grows depending on λ, following the exponential formula described in Equation 6.1 (and reproduced in this Chapter in **Section 7.3**). Since we use an exponential $ACC(RT)$, the gauge is fast at the beginning and slowing down at the end, theoretically reaching unity at infinite time. As before, participants were asked to response (by pressing a key on the keyboard) whenever they wanted to, and received a reward with probability $ACC(RT)$, or a punishment otherwise. After the response and a delay, the trial started again (the gauge is reset). In Figure 7.1 we show two examples referring to an experiment design with $\alpha = 0.25$, meaning that the gauge started a quarter full at time 0 (the beginning of the trial). The gauge represented in the figure fills up following the $ACC(RT)$ with $\lambda = 1$. On the left, we show the gauge at the beginning of the trial, whereas on the right is the gauge after one second. The top part of the figure represents exactly what is shown to the participant at each time. On the bottom part we show the corresponding $ACC(RT)$ function. The blue line in the bottom part of the figure corresponds to the $ACC(RT)$ with $\lambda = 1$, that is the one used for the gauge on the top panel. The red dashed line shows the accuracy level at that point in time. For example, after 1 second the accuracy will be around 73% (see intersection between red dashed line and blue line in the bottom-right side of the Figure 7.1). The other $ACC(RT)$ with different λ are show for comparison. For example, by using $\lambda = 0.5$, the gauge would fill up slower, and to reach the same level of accuracy it will take 2 seconds (see the intersection between the dashed red line and the orange line in the bottom-right side of the Figure). On the other hand, with $\lambda = 2.5$, the gauge will reach 73% after 0.4 seconds (intersection between red dashed line and green line in bottom-right side of the Figure).

In the version presented in the previous Chapter, participants reported to be especially annoyed by negative sound feedback. We believe that this would result in an unwanted increase of subjective punishment, such that participants would become more loss-aversive. We desired participants to be affected mostly by the punishment and reward in terms of points, and thus we eliminated, for this and the following EXACT studies, any sound feedback altogether.

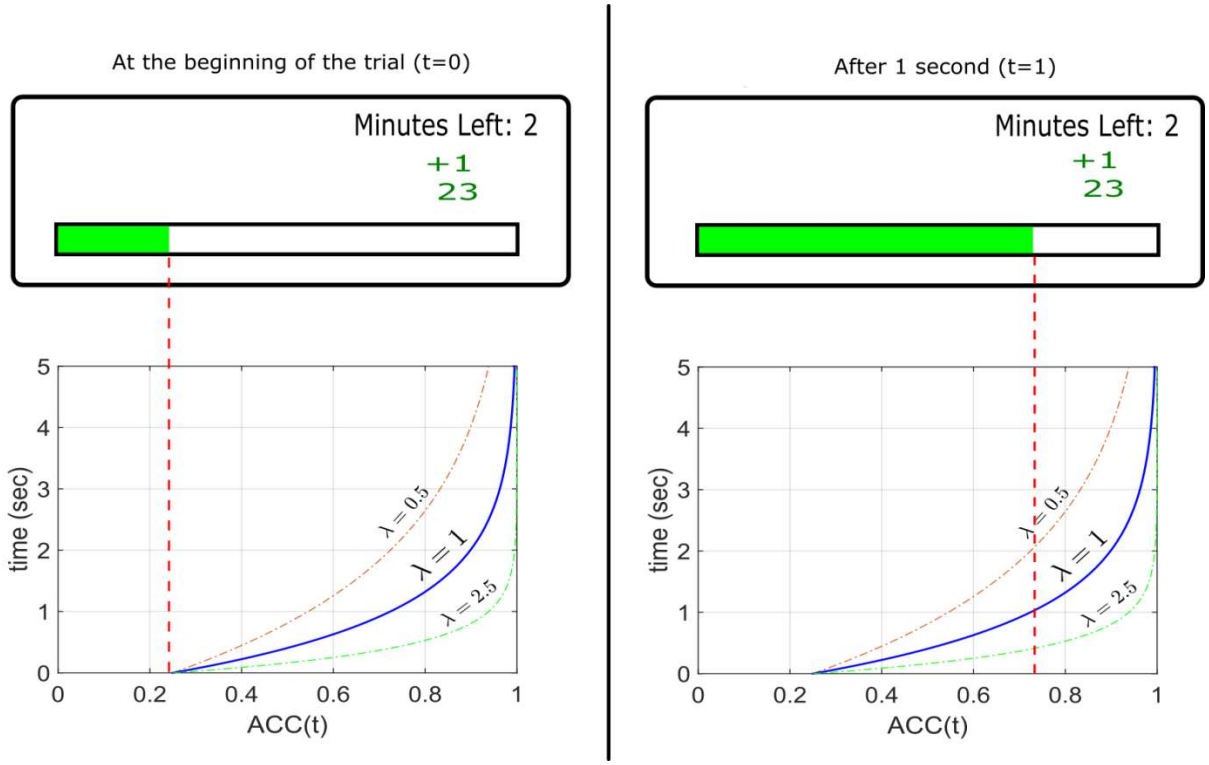


Figure 7.1

(Top) Each square represents what a participant see on the monitor during two different type of the experiment: at the beginning of the trial **(left panel)** and after 1 second **(right panel)**. The green gauge corresponds to the value of the speed-accuracy function, $ACC(RT)$, indicated in Equation 6.1 and plotted on the **bottom panels**. See text for details.

Note that the mathematical formulation of the decision strategies and the predictions discussed in Chapter 6 are exactly the same for this design (**Section 6.1.3, 6.1.4, 6.1.5**).

Another advantage of the EXACT Paradigm is the possibility of analysing the accuracy obtained for each response directly (which corresponds to the positional response along the gauge), and thus making possible analysing the distribution of accuracy. In fact, in Classic RT task, analysing the accuracy may be extremely troublesome, as hundreds of responses are needed in order to obtain a single value of accuracy for each condition. Other techniques group accuracy for each response bin within each condition (Luce, 1986), but they are rarely used as requires a very high sample for each response bin, which is often difficult to have. With the EXACT Paradigm, we can obtain accuracy directly from each the response.

7.1.1 Organization of the Chapter

In this Chapter, we aim to investigate thoroughly all the different facets of this new experimental paradigm, by manipulating three parameters to check what is the results on participants' behaviour and what decision rule can best account for participants' responses. As we saw in **Section 3.6**, each decision rule has at most 4 parameters: λ, D, α, q. In this Chapter we investigated the prediction of manipulating D and α. To increase the number of conditions in each experiment (as this allows us a more precise fitting in terms of theoretical behaviour), for each experiment we also manipulated the level of λ, corresponding to different steepness of the $ACC(RT)$ function (and thus speed of the gauge, see Figure 6.1 and 7.1). In the next Chapter, we investigate the effect of manipulating the "emphasis on accuracy" parameter q, and compare those results with a Classic RT task with a comparable payoff matrix.

In the first Experiment of this Chapter, we manipulated the D parameter by increasing the average delay across trial (D conditions). This will test the same prediction as in the previous Chapters, that is: increasing D will correspond to an increase in RT. We focus again on this prediction because it is a counterintuitive prediction that would provide a strong support to the idea that participants were optimizing their response based on the total trial time. We have already seen in previous Chapters how this idea does not seem to hold in Classic RT (and in the previous version of the EXACT Paradigm). We investigated this idea in this Chapter as well, with a more complex design that allowed us a quantitative comparison of decision rules.

In the second experiment, we investigated the most interesting prediction generated by the RR_m decision rule, that is the possibility of observing Piéron's and non-Piéron's shape of responses depending on the combination of α and q values (see **Section 2.10** and **Section 6.1.5**). In particular, we manipulated the α parameter, that is the starting point of the gauge (in a classic RT experiment this corresponds to manipulating the number of alternatives), by using 2 values α level combined with 4 λ levels, and keeping q fixed. The similar relationship is investigated in Chapter 8, and then compared with a Classic RT task.

EXPERIMENT 7.1

7.2 Stimulus Intensity, Delay, and Piéron's Law

In terms of optimal response, we have already shown that by increasing D we expect an increasing in DT^* (now defined as RT^*). For different λ, the shape of increase in shown in Figure 7.2 for RR_m (see Figure 6.2 for all decision rules). We refer to each one of the curve connecting the same D condition across different λ condition as the RT^* curve. In Figure 7.2 the RT^* is calculated by computing the optimum response for RR_m. Similarly, the result for other decision strategy is simply to increase the RT^*.

As we discussed before, the parameter λ establish how fast the speed-accuracy function increases in time and can thus be considered as the "strength" of the stimulus. This correspondence make it possible to compare the RT^* curve with the Piéron's Law function. As we explained in **Section 2.10**, Piéron's Law can be fitted for any decision strategies, if the initial fitting point is selected to be higher than the maximum point of the RT^* curve (Figure 2.6). This leads to the question of what set of parameters of the Piéron's function is affected when varying the parameter D. To answer this question, we ran several simulations. We used a similar approach to the one we are going to use for real data: we used the RT^* corresponding to 10 equally spaced values for each value of D, and we fitted the Piéron's function for each function, considering λ equivalent to stimulus strength ($RT = \gamma + k\lambda^{-\beta}$). We assumed that k, β or γ (or any combinations of those, plus a condition where no parameter changed across trials) were free to vary across D conditions, resulting in 8 different models. The best fitting (lower BIC) was given by the version with k and β free to vary across experimental condition. The goodness of fit for each curve was close to perfect ($R \cong 0.999$). The estimated α values where increasing (from 0.4 to 2) and the estimated β values where decreasing (from 0.697 to 0.673) with increase in D. The estimated γ was 0.083. This result is particularly interesting: by numerical simulation, it is clear that the RT^* for each decision strategies will tend to 0 as λ tend to infinity. However, by estimating Piéron's Law only up to a certain value of λ, the fitting procedure will not be able to capture

this feature (which, if correctly captured, would return $\gamma = 0$), returning instead a value of γ that is similar to what normally obtained in the literature (e.g. Pins & Bonnet, 1996; Bonnet et al. ,1999; Mansfield, 1973).

7.3 Methods

20 participants (8 males, 12 females) took part in the experiment. The experimental setup was almost identical to the one presented in the previous Chapter, with the major difference being that the gauge now represented $ACC(RT)$, and no sound feedback was present. Participants were shown a progressing gauge starting at some point (defined by the parameter α) and increasing in time according to $ACC(RT) = 1 + (\alpha - 1)\exp(-\lambda t)$ (presented initially in Equation 6.1, **Section 6.1**). On the top left of the screen some information were given to the participants: the amount of point earned/lost in the previous trial, the total score for the current session, and the time left in the session in minutes. If the previous trial resulted in a loss, the points were displayed in red, otherwise in green (see top part of Figure 7.1: in the black bold square we show exactly what the participants saw on their monitor, excluding the red dashed line). There was no sound feedback. Each participant started with 25 points at the beginning of each condition.

In the previous Chapter, care was needed to ensure that participants believed that there was no secret pattern or a particular schedule in according to which the reward and punishment were given. We used a training session similar to a Simple RT task, and then a training session with what we suggested to be an invisible stimulus. Here this was not necessary, as participants clearly saw the gauge increasing. We did use a practice session to get them accustomed with the task. In the practice session, after the beginning of the trial, participants were asked to press the CTRL key to respond, whereupon a reward was given with probability $ACC(RT)$, otherwise a punishment was given. In this practice session, we used punishment and reward equal to 1 ($q = 1$), a delay equal to 1 seconds, $\alpha = 0.5$ (gauge starting half full at the beginning of the trial) and $\lambda = 1$. After the response and the delay, the gauge was reset. The practice session lasted 1 minute.

After the practice session, the real experiment started. We used three D conditions, with a delay of 0.1s, 1s and 2.5seconds, and 3 λ conditions: 0.5, 1 and 5, whereas the parameter α was fixed to 0.5 (bar half full at the beginning of the trial). The order of presentation of D was randomized. Once the D condition was randomly chosen, all the λ conditions were performed (in random order) for that D, before proceeding to the next D. At the beginning of each D session, there was a 30 seconds practice session (with $\lambda = 1$) so that participant could get accustomed with the delay across trial. Each session lasted 3 minutes, with a total experimental time of about 30 minutes. Participants were informed that at the end of the session points were exchanged for money, at the rate of 1 pence for each 2 points.

Analysis. We used the Weighted Minimum Least Square method (see **Section 3.8**) to fit the four decision rules (RR, RR_m, RA and BR). Note that, compared to a normal estimation routine in a Classic RT Paradigm (for example, when estimating the parameter of a DDM, as we have done in Chapter 4 and 5), here we knew exactly what were the real experimental parameters. However, it is conceivable that participants based their responses on estimated values which may have been close but not exactly the same to the real experimental parameters. In fitting the decision rules, we used several variants by assuming that different combinations of parameters were being estimated by the participants. Therefore, we distinguish between the real experimental parameters (α, q, λ, d) and the estimated ones, which we indicate with a hat ($\hat{\alpha}, \hat{q}, \hat{\lambda}, \hat{d}$). The estimated parameters for a particular condition are indicated by using a subscript. For example, $\hat{\alpha}_{\alpha=0.5}$ is the estimated $\hat{\alpha}$ parameter for the $\alpha = 0.5$ condition. This parameter is expected to be close to α. When no subscript is used, it means that the parameter is the same across all experimental conditions.

In order to take into account discrepancy between the experimental parameter set θ and the subjective estimated $\hat{\theta}$ we propose to use a loss function in which an additional component is added to the Weighted Least Square Error. This component penalizes for the distance between real and subjective experimental parameters by using a Normal Kernel:

$$modified\ WLS = \sum_{i=1}^{n} \frac{\left(y_i - f\left(x_i|\hat{\theta}\right)\right)^2}{\sigma} + \frac{\omega}{N}\sum_{j\in S}^{N} 1 - exp\left(-\frac{\left(\widehat{\theta_j} - \theta_j\right)^2}{2\sigma_{\theta_j}^2}\right)$$

where n is the number of observation (for a single condition), N is the number of parameters for a single condition, $\hat{\theta}$ refers to the set of subjective parameters, θ to the set of experimental parameter, and ω is the weight given to the deviation from real parameter (set to 1 for this experiment). The second part of the equation is applied to the set S of parameters θ. In this experiment, we calculate the deviation for all parameters, but in the next Chapter we will calculate the deviation only on a subset. We call this measure modified WLS (mWLS). Note that σ here refers to the standard deviation of the observation of an experimental condition, whereas σ_{θ_j} refers to the standard deviation of the kernel for each θ parameter. In particular, we used a standard deviation of 3 for λ, 1 for α, 3 for D and 5 for q. These values may be deemed to be quite permissive. The reason is that, by trying several combinations of these parameters, we noted that the parameters converged to approximately the same results, given that the standard deviation values were high enough to obtain a mWLS value higher than 0 (otherwise numerical limitations of the $fitminsearch$ function would not have led to good parameter estimation). As before, we selected the best variants by considering the number of parameters by using the BIC value (**Section 3.8**).

For the analysis of accuracy distributions, analysing individual distribution was not feasible, as in some conditions the sample size was too small. Aggregating distributions was a possible solution but complicated by the fact that accuracy distributions are truncated from α to 1. The approach adopted was similar to the one used by Harris et al., 2014, for the assumed truncated Normal Distribution. The idea is to exclude distributions that show severe truncation, and aggregate the others by standardising them. We defined distributions with severe truncation as those whose two standard deviations were not inside the truncation limits of α and 1.

Note about terminology. We distinguish between the condition name, the parameter name, and the subjective parameter estimation: in this first experiment, we use different delay condition, which

we call D conditions. In terms of decision rules, they affect the D parameter. We assume that participants have their own subjective parameter estimation (including subjective estimation for the parameter D), which we call $\widehat{D}$, $\hat{\alpha}$ and so on.

7.4 Results

The optimum average score calculated assuming that participants were trying to earn as many points as possible, compared with the empirical score, is shown in Figure 7.2. Participants score (open circles) seem close to optimum average score (closed circle), except for the $D = 0.1s$ condition, especially with higher λ. The reason for this difference may be due to the fact that, for this condition, the optimum theoretical response time was quite fast (~0.172seconds), and participants settled for a slower RT. Until now we have never assumed that participant had a cost in responding related to how fast the response was, and this has been only rarely investigated (Drugowitsch et al., 2012). However, it seems reasonable that participants may decide to earn less in exchange to have a more relaxed response time, even though is difficult to be more precise than that as we do not know the relationship between speed and cost. This will be further discussed in **Section 7.5** and in the General Discussion in Chapter 9.

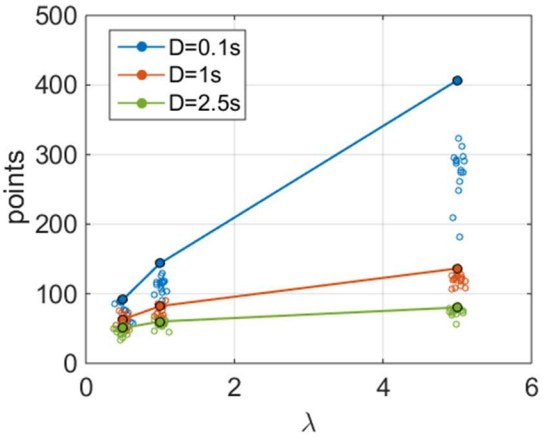


Figure 7.2

The empty circles represent the points obtained for each participant for each condition. Different D conditions correspond to different colour. The filled circles represent the expected average score with optimum performance (calculated with RR_m), with a line connecting the same D conditions. Noise on the horizontal axes is added to increase readability. Participants

appeared to be responsive to the changing of the condition, but they were consistently sub-optimal, especially for the higher λ and the $D = 0.1s$ condition. This is probably due to the fact that an optimum score can be only obtained by generating very fast responses.

7.4.1 Response Signal and Response to Feedback

As before, we analysed the stability of responses. In the previous experiment, we saw that the average number and magnitude of shift index for each condition was much higher than in a Classic RT design (Figure 6.5 and 6.6). We hypothesised that the reason was to be found in the fact that participants did not have access to $ACC(RT)$ and needed to estimate it, forcing them to go through a more intense explorative phase. The number of shifts was, in fact, remarkably reduced in this experiment (compare, for example, the $D = 0s$ condition of the previous experiment with the $D = 0.1s$ condition of the current one). However, these values are still much higher than those showed in Classic RT task (which contains usually around one shifting per condition, as shown in Figure 6.6). The number of shifts averaged across participants for each condition is shown in Table 7.1. An example from a typical participant is shown in Figure 7.3.

Table 7.1 Average shift index for each condition

	$\lambda = 0.5$	$\lambda = 1$	$\lambda = 5$
$D = 0.1s$	4.4375	5.125	4.8125
$D = 1s$	2.875	3.1875	2.4375
$D = 2.5s$	1.5	1.625	1.5

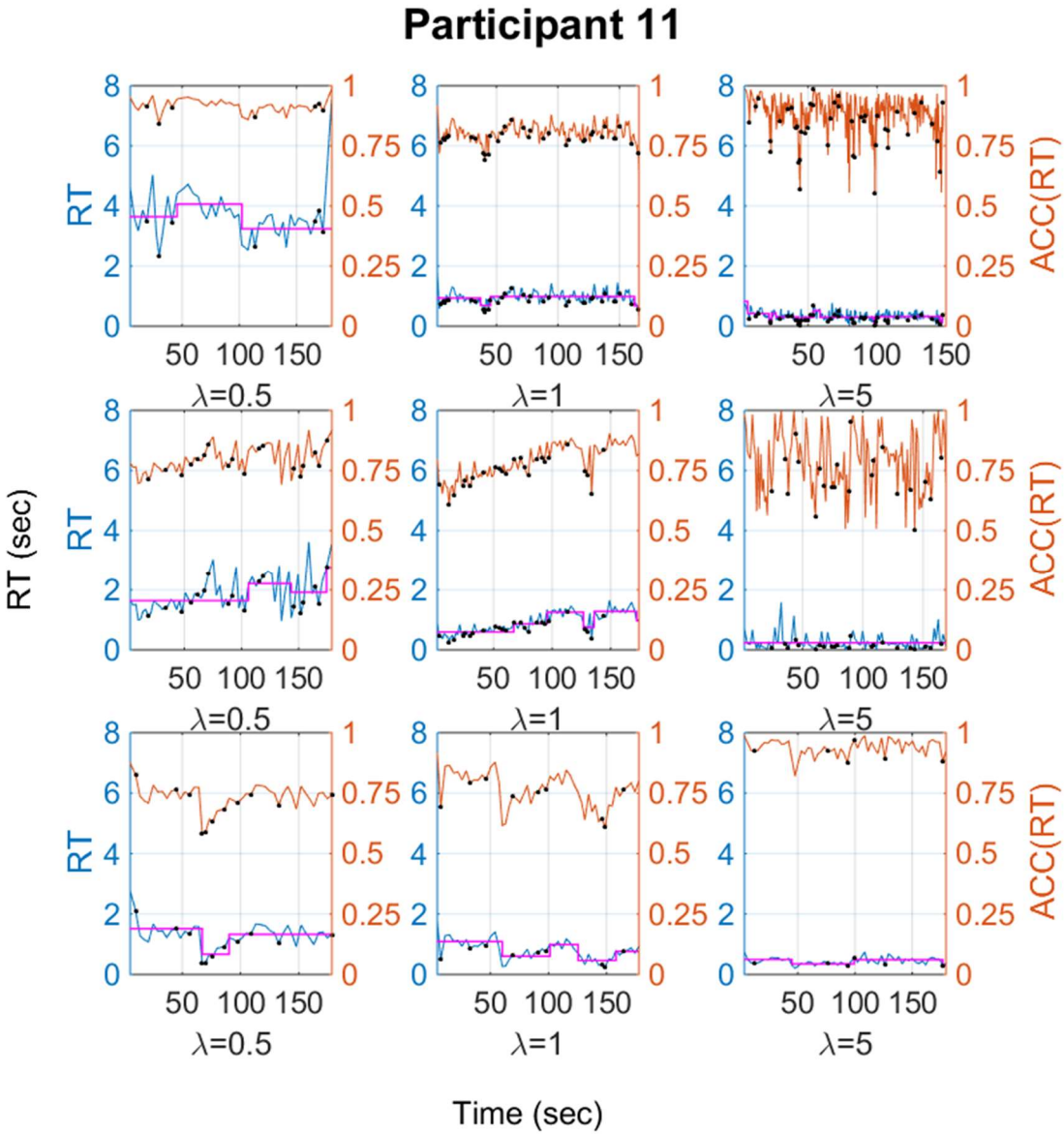


Figure 7.3

Response Signal for a typical individual, for each condition. Each row corresponds to a different D condition (from the top to the bottom, $D = 0.1s, D = 1s, D = 2.5s$). The blue lines represent the response in terms of RT (left vertical axes), the orange lines represent the response in terms of $ACC(RT)$ (right vertical axes, see Equation 6.1 for the relationship between RT and $ACC(RT)$). The blue dots on the lines correspond to error responses. The magenta lines are the result of STARS method, and they represent the detected shifting in strategy. With this version of the EXACT Paradigm participants clearly produces responses that are more stable and have less shifting (compare with Figure 6.5).

We used participants' shift index values for each condition as an input to a repeated measure ANOVA, finding that only the D conditions significantly affected the amount of shifting (for RSI: $F_{2,30} = 46.512, p < 0.0001$, for λ: $F_{2,30} = 6.5, p = 0.06$). If participant already know $ACC(RT)$ because is shown to them, what is the reason for the change of strategy? It is possible that, even though participants know $ACC(RT)$, they still need to proceed by trial and error to find the optimum that maximizes their decision rule. The reason why it appeared to depend on D (with some non-significant dependency on λ, at least for the $D = 0.1s$ condition), may be that with low D participants have the chance to perform many more trials, and therefore they are more willing to use some of these trials for exploring different strategies. Note that we also measured the magnitude of the shift, which did not appear to be affected by different conditions (average shift magnitude across all condition was equal to 0.79).

Despite this relatively high shifting values, the response signal appeared to be fairly stable: by aggregating signal responses in bins and calculating the correlation across trial length for mean and standard deviation of the response (similar to what done in the previous experiment), we did not find any significant trend in any of the conditions. As noted before, the lack of changing of standard deviation implies that participants were not adjusting the precision of their responses, at least not during the few minutes given to them.

We did not find any relevant difference in the way participants adjusted their response after punishment or rewards. On average, participants slowed down of about 0.3seconds after a punishment

and speeds up of about 0.02 seconds after a reward, with negligible differences amongst different conditions. As before, we found a much more remarkable effect of punishment than rewards, which may be due to the fact that there were many more rewards and thus a finer adjustment was possible.

7.4.2 Effect of D and λ on RT and $ACC(RT)$

The results in terms of mdRT and accuracy of response $ACC(RT)$ are shown in Figure 7.4 for the six experimental conditions (lines in the Figure connects different λ conditions across the same D condition). The optimum response (calculated with RR_m) for each different condition is also plotted in the figure (circles with black edges). The plot in terms of $ACC(RT)$ is equivalent to showing at what point on the bar, on average, participant responded. This plot tells us something important about the paradigm, that is: participants were not merely pointing to the same spatial point on the bar, but they were actually changing optimum point based on experimental parameter. In fact, due to different speed of the bar for different λ, if participant were merely aiming at the same accuracy (that is, at the same position on the gauge) for different λ conditions, we would have still observed a decreasing shape in RT for higher λ, but we would have also observed a flat line on $ACC(RT)$. In our dataset, however, we observed an increasing trend in $ACC(RT)$ which confirmed that participants were adjusting their strategy according to experimental conditions.

As expected, changing λ had a clear effect of changing the mdRT ($F_{2,30} = 38.842, p < 0.0001$). Even though changing D appeared to slightly increase mdRT, this effect did not reach significance ($F_{2,30} = 1.902, p = 0.167$). The effect on accuracy of response mimicked the result for median rate, in that only λ had a significant effect on accuracy (λ: $F_{2,30} = 26.99, p < 0.0001$ and D: $F_{2,30} = 2.5, p = 0.099$). Post hoc pairwise comparisons with Bonferroni correction revealed that all the three λ conditions were significantly different from each other at $p < 0.0001$, whereas none of the D conditions were significantly different from each other. Segregating participant according to their performance did not appear to make any relevant difference. It was also clear that participants' RT were far away from optimum response, as they were all very similar across the three D conditions, being faster for the long D condition and slower for the fast D condition.

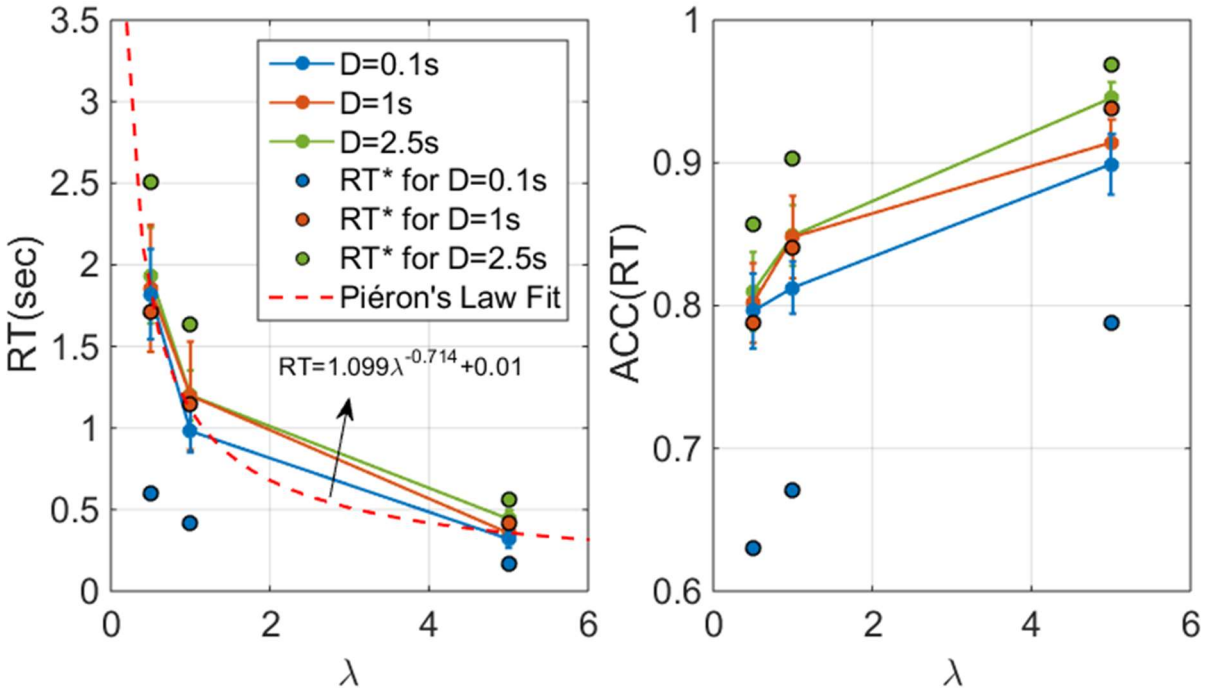


Figure 7.4

The **right panel** filled circles indicate the median of median of responses (mdRT) across individuals for each D and λ condition. The bars on the circles correspond to 1 standard error. The mdRT within the same D conditions are connected with a line. The circles with the black edges correspond to the optimum response for each condition. The red dashed line corresponds to a the best fitting of a Piéron's Law. In this case, only one line was used as the best model resulted to be the one without any variation in parameters across different D conditions. The **right panel** shows the corresponding values in terms of median of median $ACC(R$.

We also analysed the results in terms of individual participants. Whereas the λ factor was very consistent across different participants, the D effect was more variable, sometime producing slower responses, sometime faster, sometime unchanged. This result is consistent with results from the previous Chapters: the delay across trial has a negligible or null effect (however, see **Section 7.5** for other possible explanations).

We fitted the aggregated results to Piéron's Law. Different models were generated by assuming k, β or γ (or any combination of those) free to vary across D conditions, and compared them with the modified BIC for mean least squares which takes into account the number of parameters. The best fitting model was the one with no parameter free across D conditions, which confirmed the small effect obtained by different D conditions (red dashed line in Figure 7.4). Similarly to what obtained with

simulations (see **Section 7.2**), we found a γ value equal to 0.01 (Table 7.3). The main point of this fitting procedure was to confirm that a Piéron's shape can be replicated in the EXACT Paradigm (as predicted), which lead the possibility that Piéron's Law is not an artefact of the perceptual process (as sometime argued, see Stafford & Gurney, 2004) but a result of a simple decision rule strategy with a monotonically increasing $ACC(RT)$. Moreover, the fact that we were able to reproduce a Piéron's Law confirmed the similarities between this experiment and a Classic RT task.

Table 7.2 Piéron's Law Model

Model	BIC	df
No free	**-5.58142**	**3**
k free	-4.69314	5
β free	-3.06478	5
γ free	-4.62256	5
k and β free	-4.38543	7
k and γ free	-5.12028	7
β and γ free	-1.60171	7
k, β and γ free	-2.3892	9

Table 7.3 Parameters for the best model

Model 1	
BIC	-5.58142
WLS	0.0048
$\boldsymbol{k}$	1.099
$\boldsymbol{\beta}$	0.714
$\boldsymbol{\gamma}$	0.01

7.4.3 Fitting Decision Rules

We fitted the Decision Strategy by taking into account the distance between the estimated parameters and the experimental ones (see Methods, **Section 7.3**). Each one of the four decision strategies (RR, RR_m, RA and BR) was evaluated with two different versions: in one, the estimated parameter $\widehat{D}$ was allowed to vary across different delay conditions (D); in the other version, $\widehat{D}$ was fixed across the two conditions. In this way, we could test whether the same parameter $\widehat{D}$ was used for different D conditions, as other results seems to indicate. Note that for BR we used only one variant, as this decision strategy did not use the parameter D at all. The results agreed with the previous Chapters' result: the variants with the lowest BIC was RR_m with fixed $\widehat{D}$ across experimental conditions (Table 7.3 and Figure 7.5). Note that the RR_m with free $\widehat{D}$ for each D condition is the second to best model, and in fact it obtained the lowest modified WLS value but it was penalized for the number of parameters

(8 against 6 of the other variants). The worst model was RR with free $\widehat{D}$ parameter. Table 7.4 shows the BIC and modified WLS value for each model and variants, whereas Tables 7.5 shows the resulting parameters for the best model, the RR_m variants with fixed $\widehat{D}$.

Table 7.4 Results for the decision rules fitting

	BIC	df
$\boldsymbol{RR_m\ (free\ \widehat{D})}$	-1.8647	8
$\boldsymbol{RR_m\ (fixed\ \widehat{D})}$	**-1.95996**	**6**
$\boldsymbol{RA\ (free\ \widehat{D})}$	-1.72593	8
$\boldsymbol{RA\ (fixed\ \widehat{D})}$	-1.69524	6
$\boldsymbol{BR}$	1.785779	5
$\boldsymbol{RR\ (free\ \widehat{D})}$	1.919916	7
$\boldsymbol{RR\ (fixed\ \widehat{D})}$	0.233448	5

Table 7.5 Best models' parameters (ΣmWLS is the sum of the modified WLS across conditions)

$\boldsymbol{RR_m(free\ \widehat{D})}$		$\boldsymbol{RR_m(fixed\ \widehat{D})}$	
BIC	**-1.865**	**BIC**	-1.96
ΣmWLS	**0.1549**	**ΣmWLS**	0.22954
$\boldsymbol{\hat{\lambda}_{\lambda=0.5}}$	0.613	$\boldsymbol{\hat{\lambda}_{\lambda=0.5}}$	0.438
$\boldsymbol{\hat{\lambda}_{\lambda=1}}$	1.156	$\boldsymbol{\hat{\lambda}_{\lambda=1}}$	0.829
$\boldsymbol{\hat{\lambda}_{\lambda=5}}$	5.556	$\boldsymbol{\hat{\lambda}_{\lambda=5}}$	4.73
$\boldsymbol{\hat{\alpha}}$	0.485	$\boldsymbol{\hat{\alpha}}$	0.491
$\boldsymbol{\widehat{D}_{D=0.1}}$	0.372	$\boldsymbol{\widehat{D}}$	0.384
$\boldsymbol{\widehat{D}_{D=1}}$	0.632	$\boldsymbol{\hat{q}}$	1.177
$\boldsymbol{\widehat{D}_{D=2.5}}$	0.652		
$\boldsymbol{\hat{q}}$	1.337		

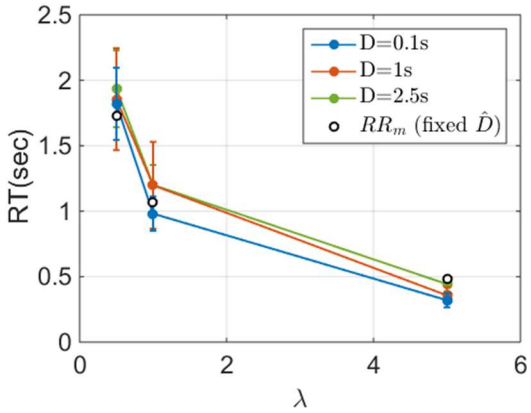


Figure 7.5

Median of median RT, the same as plotted in Figure 7.5. The white circles show the expected values with the RR_m model with estimated parameters (see Table 7.5)

It is interesting to note that, in both cases, the q parameter is higher than 1, meaning that participants put much more emphasis on accuracy than on speed. In terms of RR_m, this means that the subjective value of punishment was estimated to be higher than the subjective value of the reward: in other words, participants were generally loss-aversive. Furthermore, the estimated $\widehat{D}$ values for the RR_m with free $\widehat{D}$ were, as expected, all very close to each other. This confirmed what we found by running the ANOVA on the dataset: the difference between D conditions is small or non-existent. Note how, in this case, a qualitative prediction (the first level of analysis we discussed in **Section 2.7**), would have led in believing that the BR model was indeed a viable alternative, has it predicts no effect with varying D conditions. Instead, fitting the model revealed that the BR fitting is indeed a poor model for this task.

7.4.4 Distribution Analysis

The Vincentized distributions in time domain for each λ and D condition are shown in Figure 7.6. Mean, standard deviation and skewness were calculated by averaging these statistics across all participants. Distributions appeared right skewed, as usually found in classic RT experiments, but the skewness is much lower than what found in Classic RT (for example, compare with Figure 4.6 or Figure 5.5). We found a very strong effect of λ (but not D) on standard deviation (for λ: $F_{2,30} = 4.594, p \cong 0$, for D: $F_{2,30} = 0.605, p = 0.552$), and a non-significant effect on skewness for both λ and D (for λ: $F_{2,30} = 0.512, p = 0.605$, for D: $F_{2,30} = 0.813, p = 0.453$). Consistently with what found in the

previous Chapter, the distributions in the rate domain did not appear close to a Normal distribution, which was confirmed by a Kolmogorov-Smirnov test which rejected normality for all distributions

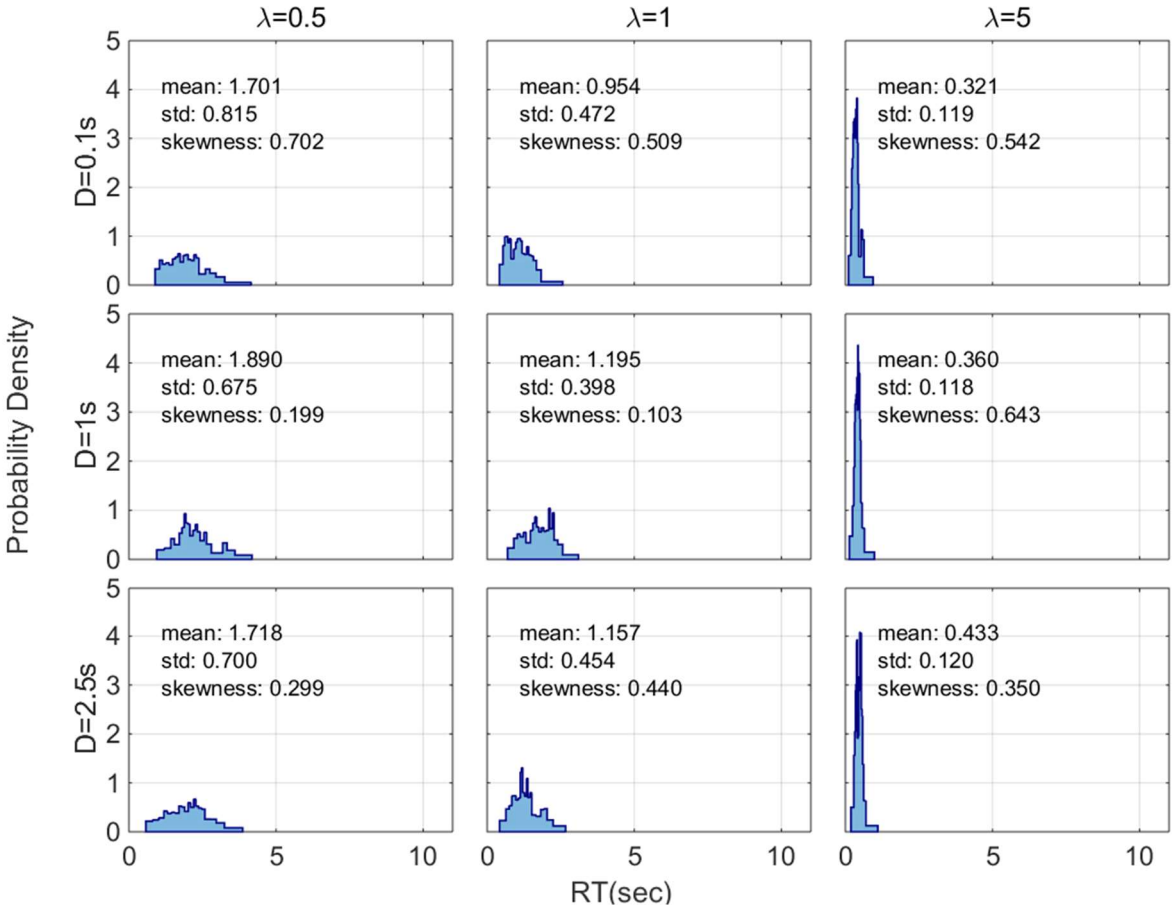


Figure 7.6

Vincentized RT distributions for each condition. Increasing λ corresponds to a steep decrease on standard deviation.

An interesting point of this paradigm regards the possibility of analysing the accuracy distributions, which corresponds to the positional response along the gauge. The $ACC(RT)$ distributions, after standardising (see Method in this Chapter) and aggregating across participants are shown in Figure 7.4. Cutting the severely truncated distributions, however, did not produce a relevant difference (blue lines correspond to the original dataset, yellow line to the modified one). We ran a Kolmogorov-Smirnov test for normality, and it resulted that 5 distributions out of 9 were not significantly different to a Normal Distribution (at statistical level of $\alpha = 0.01$). Note that for the Distribution that most diverged from normality (with $D = 1s$ and $\lambda = 5$) we also found the most severe truncation value (12 datasets were

excluded for that condition). It is however difficult to conclude anything with confidence with this dataset, so we postpone a more detailed discussion to Section 7.10.

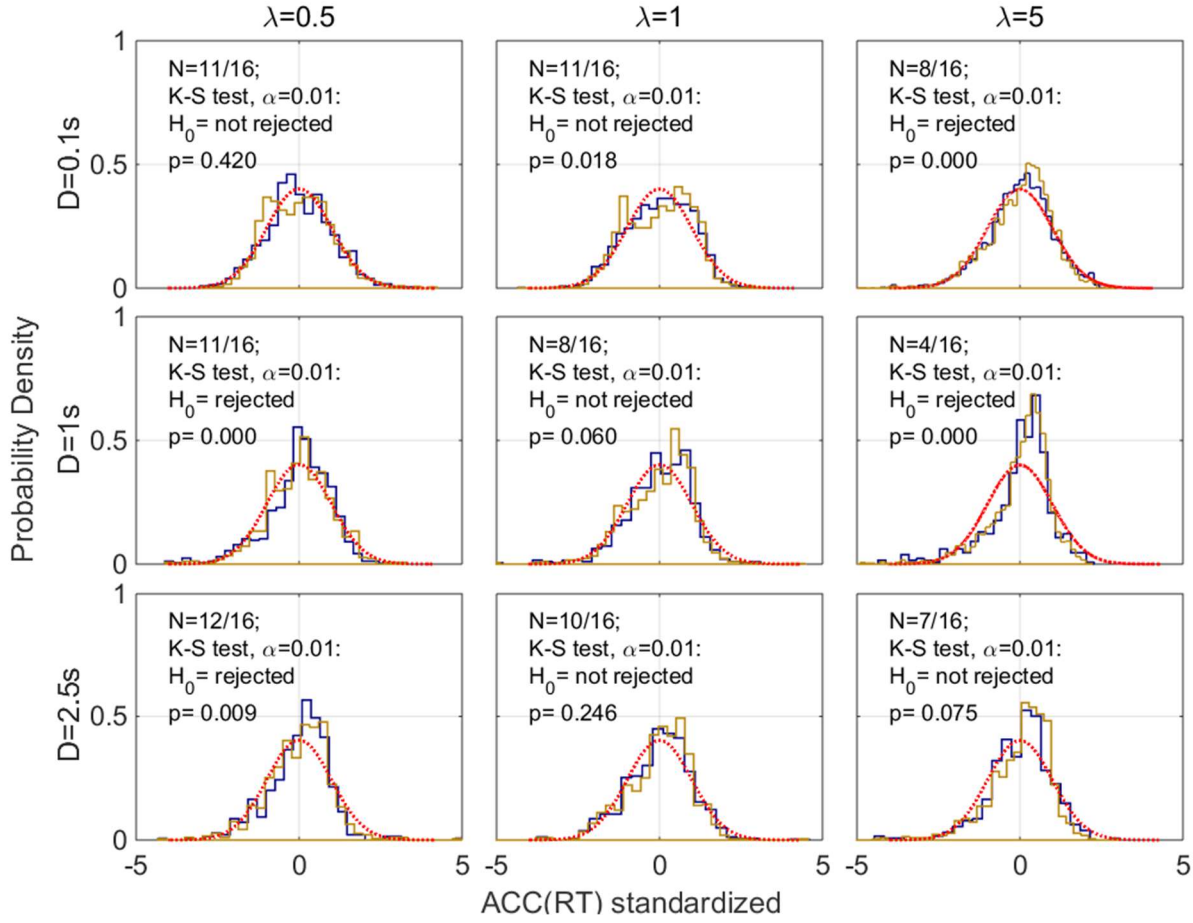


Figure 7.7

Aggregated and standardised accuracy distributions, before and after excluding the severely truncated distributions (blue lines show the original dataset, yellow line the modified one, and the red dashed line corresponds to a Standard Normal Distribution). On each panel, the number of excluded distributions is indicated (N), and the result of a Kolmogorov-Smirnov test for normality.

7.5 Discussion

Using this complete version of the EXACT Paradigm resulted in having participant perform a simple task in which they had complete information about the $ACC(R$ at each time step, obtaining therefore more reliable measures of behaviour. This also resulted in more stable response signals, with less shifting in strategy and no relevant trend across conditions. Moreover, using a more complicated design with two factors (λ and D) allowed us to fit more precisely the decision strategies to participants'

RT. Altogether, these results hinted to the fact that participants are spending less time in estimating the correct $ACC(RT)$ and more time in finding the best strategy to maximize the decision function.

In this study, we also introduced a new way of analysing the results by computing the accuracy distribution. With the EXACT Paradigm, we have simultaneous access to RT and $ACC(RT)$, and we can therefore make theoretical prediction about the $ACC(RT)$ distribution. There is, in fact, no reason but practical ones about why we should base most of our analysis on the RT distribution. In fact, $ACC(RT)$ appear approximately Normal distributed (at least for those conditions in which not too many datasets were cut due to sever truncation), which may pinpoint to the important fact that participants may aim to some special value of accuracy, and RT are a simple by-product of this computation. Even though in this experiment the analysis of $ACC(RT)$ did not lead to any univocal conclusion, it opens up the possibility of analysis in new directions that are not possible in a Classic RT (see also **Section 7.10**).

This experiment confirmed the results of the previous Chapter: participants do not appear to vary their response depending on the total trial time D. Despite this result, the best strategy still resulted to be RR_m, which actually does take into account a parameter D. What are the possible explanations for that?

1) Participants were poorly estimating $\widehat{D}$. This is the interpretation given by other authors (Zacksenhouse, Bogacz, & Holmes, 2009), but it seems hardly applicable in this case, where the estimated $\widehat{D}$ is 0.384 and the real one is as high as 2.5seconds. Furthermore, the experiments in previous Chapters appear to disconfirm this hypothesis.

2) Participants were not using the delay across trial value at all, and the parameter D within RR_m should be interpreted as related to some other experimental factor. This is important, as in this case RR_m lose its meaning of being the optimal strategy, but become a unknown variable that participants were using to take a decision.

3) Participants were using and estimating the D parameter, but the section of the trial they were using to do so did not correspond to the whole trial duration: they were using a different "epoch" for maximizing their response (see General Discussion in Chapter 9)

4) Participants were using a standard value for D, and changing it only when it was convenient. In our experiment, by using the real D condition as the $\widehat{D}$ parameter, the total expected score with optimum RT^* was calculated to be equal to ~804 points (summed across all 6 conditions). By using the estimated $\widehat{D}$ values for the free $\widehat{D}$ and RR_m strategy (0.37s, 0.63s and 0.65s) the total score was equal to 791. This difference is not extremely large and, for our experiment, participants may have decided that it was not worth to compute different optimum responses for different delays when the difference in terms of expected reward was not so high. Unfortunately, this possibility was not considered beforehand, as an experiment with high gross differences with different D conditions may have been designed to rule out or confirm this explanation.

EXPERIMENT 7.2

7.6 Stimulus Intensity and Starting Point

In this experiment, we investigated the effect of different α values on responses, by also varying the difficult of the stimulus λ in order to obtain different RT^* curves. In terms of EXACT Paradigm, λ represents the rate of accumulating accuracy in time, and α represents the accuracy at the beginning of the trial. In a Classic RT task, λ is related with the difficulty of the task, whereas the interpretation of α is more problematic. This could in fact be interpreted as the number of response alternatives (that is, $N = 1/\alpha$) or as the probability of each stimulus (in a multiple-choice task) to be the correct one. In Classic RT paradigm, these concepts are related, as by increasing the number of alternatives we are automatically decreasing the probability of each stimulus to be the correct. The relationship between the number of response stimuli and the corresponding RT has been analysed countless times (for a review, see Luce, 1986), the main finding corresponding to the fact that mean RT increases logarithmically as the number of alternatives increases (Hick's law, see Hick, 1952; Merkel, 1885). On

the other hand, the relationship between mean RT and stimulus probability is less clear, as previous literature usually uses too few data points to allow any fitting. The results nevertheless suggested an increasing non-linear relationship between mean RT and stimulus probability (e.g. Simen et al., 2009; Drazin, 1961). However, drawing a direct parallel between the α used in the EXACT paradigm and these other two experimental variables may be misleading. Firstly, by changing the probability of correct response in a 2AFC task, we can obtain only values from 0.5 to 1, as this values are complementary with the other stimulus probability (decreasing one will increase the other, so that values lower than 0.5 will result in values higher than 0.5 in the alternative). On the other hand, by varying the number of alternatives we can only test probability values from 0 to 0.5: values higher than 0.5 can be achieved only by having less than 2 alternatives, and is not clear how to accomplish this in a choice task. This could be accomplished by a completely different design: for example, $\alpha > 0.5$ may be obtained by using N alternatives with *at least* $N/2$ correct choices. The problem with this approach is that it seems likely to have an effect also on the rate of stimulus accumulation. For example, for $\alpha = 0.5$ we need to have 2 correct stimuli out of 4 alternatives; with $\alpha = 0.75$ we set three 3 stimuli out of 4. By changing the number of stimuli we increase the amount of information accrued at each time step, and therefore we affect the speed of increasing accuracy in time. Therefore, in a Classic RT task we would need different designs to test α from the range of 0 to 1, and this would have slightly different interpretation too. Instead, with the EXACT Paradigm we can more freely change the α parameter within the 0-1 range, by using the same design, without affecting the rate of accumulation of information.

One of the main aims of this experiment was testing the prediction of the RR_m decision rule of obtaining Piéron's and non-Piéron's shape by varying the starting point (α) of the gauge (Figure 7.7, see also **Section 2.10** and **Section 6.1.5**). In particular, we used two α values (0.25 and 0.75) coupled with 4 λ conditions in order to obtain a RT^* curve for each α condition. Note that, to the best of our knowledge, the peculiar non-Piéron's shape has not been observed in the past. Recall from **Section 6.1.5** that a non-Piéron's shape is obtained whenever $\alpha > \frac{q}{q+1}$. By setting the emphasis on accuracy parameter q equal to 1, we expected to obtain a non-Piéron's shape for $\alpha > 0.5$. However, we have

noted in the previous experiments that participants tended to set their subjective q value higher than the experimental one (possibly because they are loss-aversive). Therefore, in order to observe the non-Piéron's shape, we adjusted α accordingly. We opted for using $\alpha = 0.75$ and, symmetrically, $\alpha = 0.25$ for the condition that predicted the Piéron's shape. As we will see, if participant were perfectly estimating all the parameters and were trying to maximize the amount of points earned (by having $q = 1$) this would result in RT^* always equal to 0 for all λ conditions when $\alpha = 0.75$, which may be physically unreachable by participants. We do not deem this to be a problem since, as said, participants will most likely use a higher q, which would result in a slower, and less problematic, RT^*.

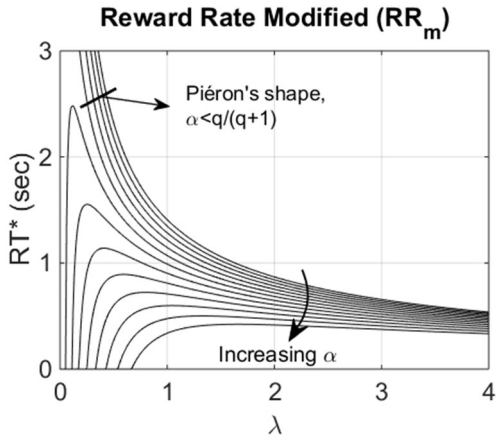


Figure 7.8

Several RT^* curves for different parameters of the RR_m decision strategies. Note how, with some parameters combination, the curves goes to infinity when λ goes to 0 (set of curves indicated by the thick black line).

This study allows us to apply the EXACT Paradigm in order to understand the effect of changing alternatives in terms of response signals, response distributions, and accuracy distributions. Furthermore, it will allow us to understand difference across several performance groups, and to distinguish between different decision strategies used by participants, both in the aggregated dataset, that in the grouped one.

7.7 Methods

We recruited 20 participants (17 females and 3 males) to take part in this experiment. The experimental setup was almost identical to the one used previously: we used a short training session ($\lambda = 1, \alpha = 0.5$) to get participants accustomed with the task. Participants saw a progressing gauge

starting at α and increasing in time according to $ACC(RT)$ presented in Equation 6.1, and they were instructed to press the CTRL whenever they wanted to request a reward. A reward of 1 point was given with probability $ACC(RT)$ (when a reward was not given, a punishment was given instead, that is 1 point was taken away). After the response, we showed a delay (white screen with a writing saying "Please Wait") and then the trial started again (the gauge was reset). On the screen, participants had information about the amount of points earned-lost in the previous trial (red if lost, green if earned) and the total amount of point for that condition (see Figure 7.1). Each participant started with 25 points. As before, points were exchanged for money (2 points=1 penny). Each participant underwent 2x4 conditions: the starting point was set to either $\alpha = 0.25$ or $\alpha = 0.75$, meaning that at the very beginning of the trial the chance of receiving a reward was equal to α and the change of receiving a punishment was equal to 1-α. We used 4 different gauge speeds $\lambda = 0.16, 0.33, 1, 2$ (higher λ corresponding to a faster increase of the gauge). Each condition lasted 3 minutes. Before the starting of each α condition, participants underwent a short training session with $\lambda = 1$ and α equal to the value of the condition. The experiment was organized so that participants performed all λ conditions for a certain α before performing the λ conditions for the other α (the order of α and λ was randomized for each participant). We used a fixed delay D across conditions of 0.5 seconds, and we did not use any timeout. To increase motivation, participants were informed that the participant who earned the most points would receive a £10 Amazon Voucher.

7.8 Results

7.8.1 Response Signal and Behavioural Responses

By comparing the individual total score for each condition (Figure 7.8) we saw that, with $\alpha = 0.25$, most of participants' total score (empty circles) systematically increased with increasing λ, and they got closer to the total expected reward. With $\alpha = 0.75$, the total expected reward (that is, the optimum response for RR_m by assuming participants are maximizing the number of point earned) was always equal to 205, as the optimum RT^* is equal to zero in this case (and the optimal accuracy is always

$ACC(0) = 0.75$). Note that an increase in participant's score did not necessarily corresponds to higher RT, as this increase may be simply due to higher λ which results in more trials.

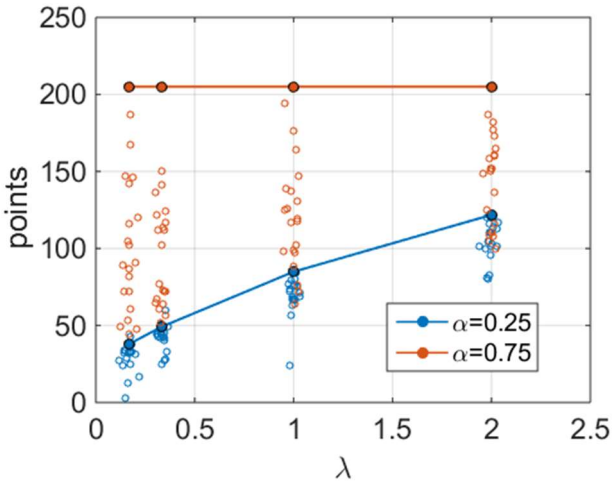


Figure 7.9

The empty circles represent the points obtained for each participant for each condition. Different D conditions correspond to different colour. The filled circles represent the expected average score with optimum performance (calculated with RR_m), with a line connecting the same α conditions. Noise on the horizontal axes is added to increase readability. As in the previous experiment, participants are consistently sub-optimal, especially with the $\alpha = 0.75$ conditions.

The analysis of response signal by using STARS method on each participant and each condition resulted in similar results as in the previous experiment. The average number of shifts index across participants are reported in Table 7.6. There was a significant difference only between different λ levels ($F_{3,57} = 16.187, p < 0.0001$), even though the values for $\alpha = 0.75$ were slightly higher (apart from the last λ). The values were consistent with the previous experiment, and the increase seemed to be due to the higher number of trials due to shorter responses with higher λ and higher α, which thus resulted in more trials and more chances for participants to change their strategy.

Table 7.6 Average shift index for each condition

	$\lambda = 0.16$	$\lambda = 0.33$	$\lambda = 1$	$\lambda = 2$
$\alpha = 0.25$	0.65	1.4	2.6	3.7
$\alpha = 0.75$	2.1	2	4.1	3.4

By aggregating response signals across participants, we could not observe any significant trend. We also calculated the change in standard deviation and, again, no significant trend was detected with this methodology. Therefore, similarly to the previous experiment, it appeared that using a visible $ACC(RT)$ mitigated the issue found in Chapter 6 about the high number of change of strategies and the presence of trends, making the results more like those obtained through a Classic RT task.

The results in terms of mdRT are shown on the left panel of Figure 7.9 (left), whereas results in terms of $ACC(RT)$ are shown on the right panel of Figure 7.9. As explained before, changes in $ACC(RT)$ for two different conditions is a good indicator that participants were actually sensible to the change of conditions, and were not just responding at the same gauge position (this could not be observed by uniquely looking at the RT responses, as responding at the same $ACC(RT)$ position would result in decreasing RT when λ is increasing). The optimum response RT^* and the optimum accuracy $ACC(RT^*)$ for each condition are represented by blue circles with black edges (for $\alpha = 0.25$) and orange circles with black edges (for $\alpha = 0.75$). Note how the optimum response was always $RT^* = 0$ for $\alpha = 0.75$, which corresponds to $ACC(RT^*) = 0.75$. The effect of both λ and α conditions on median RT was significant (for $\lambda: F_{3,57} = 42.291, p \cong 0$, for $\alpha: F_{1,19} = 45.762, p \cong 0$) and their interaction was also significant ($F_{3,57} = 17.95, p \cong 0$). Post hoc pairwise comparisons using Bonferroni correction showed that each λ condition was significantly different from each other ($p < 0.0001$). Similarly, the effect of both λ and α conditions on accuracy of response was significant ($\lambda: F_{3,57} = 27.695, p < 0.0001$, for $\alpha: F_{1,19} = 17.476, p < 0.001$). Post-hoc pairwise comparisons (by applying Bonferroni correction) showed the accuracy for each λ condition was significantly different from each other at $p < 0.0001$ (a part from the third and fourth λ condition, where $p = 0.454$).

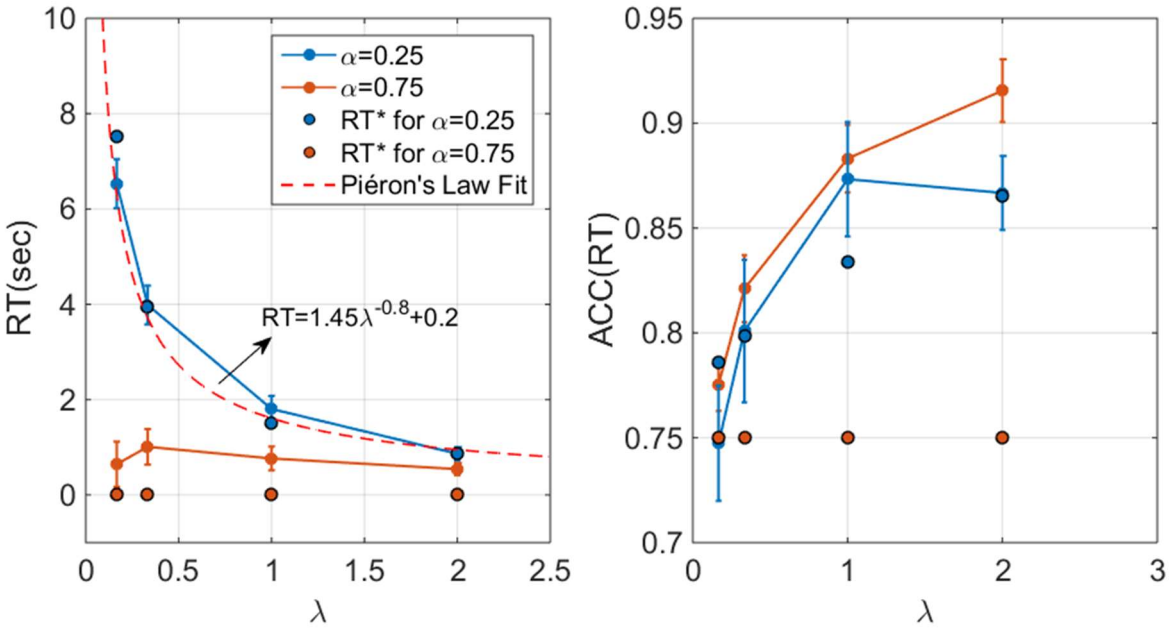


Figure 7.10

The circles indicate median of median RT (**left panel**) and median of median of $ACC(RT)$ (**right panel**). The bar indicates one standard error. The circles with black edges indicate the optimum responses, calculated with RR_m. The red dashed line corresponds to the Piéron's Law function fitted only on the $\alpha = 0.25$ condition, as the $\alpha = 0.75$ clearly does not follow Piéron's law. Note how the optimum responses for the $\alpha = 0.75$ condition are equal to 0 (orange circles on the left panel) which corresponds to an $ACC(RT)$ of 0.75 (on the right panel).

With $\alpha = 0.25$ the obtained mdRT were very close to the optimum expected reward, and they follow a Piéron's shape, as expected, which was confirmed by fitting the Piéron's function on this condition ($R\sim = 0.99$). The estimated values for the Piéron's function ($RT = \gamma + k\lambda^{-\beta}$) were $k = 1.45$, $\beta = 0.8$ and $\gamma = 0.2$. Most interesting was the peculiar shape obtained with $\alpha = 0.75$, which corresponded to the non-Piéron's shape expected with high level of α (Figure 7.7). Note that this was not an artefact of averaging across different condition, as this shape was actually observed in most participants (15 out of 20 participants). Such strategy, counterintuitive and never observed in Piéron's Law-type experiment, has been reproduced here, which implies that participants were indeed using an optimum decision strategy which can take into account different starting points.

7.8.2 Fitting Models to Aggregated Data

We fitted several variants of the decision strategies to the observation. Even though by looking at the data is obvious that the α condition had a relevant effect in the dataset, we still compared a model with free or fixed $\hat{\alpha}$ across α conditions. The results, in terms of both BIC are shown in Table 7.7. The estimated BIC values with fixed $\hat{\alpha}$ (Table 7.8) are much higher compared to the version with free $\hat{\alpha}$. We can conclude that, as expected, manipulating the starting point α had a remarkable effect on $\hat{\alpha}$ parameter.

Table 7.7 Results for the decision rules fitting (free or fixed $\hat{\alpha}$)

	BIC	df
$\boldsymbol{RR_m}$ $\boldsymbol{(free\ \hat{\alpha})}$	**-0.55746**	**8**
$\boldsymbol{RR_m}$ $\boldsymbol{(fixed\ \hat{\alpha})}$	2.356767	7
$\boldsymbol{RA}$ $\boldsymbol{(free\ \hat{\alpha})}$	-0.12927	8
$\boldsymbol{RA}$ $\boldsymbol{(fixed\ \hat{\alpha})}$	2.350711	7
$\boldsymbol{BR}$ $\boldsymbol{(free\ \hat{\alpha})}$	1.205431	7
$\boldsymbol{BR}$ $\boldsymbol{(fixed\ \hat{\alpha})}$	2.350593	6
$\boldsymbol{RR}$ $\boldsymbol{(free\ \hat{\alpha})}$	1.84737	7
$\boldsymbol{RR}$ $\boldsymbol{(fixed\ \hat{\alpha})}$	2.199422	6

Table 7.8 Best models' parameters

$RR_m(free\ \hat{\alpha})$		$RA(free\ \hat{\alpha})$	
$\hat{\lambda}_{\lambda=0.16}$	0.322	$\hat{\lambda}_{\lambda=0.16}$	0.788
$\hat{\lambda}_{\lambda=0.33}$	0.574	$\hat{\lambda}_{\lambda=0.33}$	1.219
$\hat{\lambda}_{\lambda=1}$	1.901	$\hat{\lambda}_{\lambda=1}$	2.052
$\hat{\lambda}_{\lambda=2}$	3.176	$\hat{\lambda}_{\lambda=2}$	4.087
$\hat{\alpha}_{\alpha=0.25}$	0.29	$\hat{\alpha}_{\alpha=0.25}$	0.286
$\hat{\alpha}_{\alpha=0.75}$	0.749	$\hat{\alpha}_{\alpha=0.75}$	0.824
$\hat{D}$	1.028	$\hat{D}$	0.389
$\hat{q}$	2.202	$\hat{q}$	1.964

The best model was RR_m. The estimated parameters (shown in Table 7.8) for this model were reasonably close with the real parameter. The $\hat{q}$ value was twice the experimental q, which corresponds to a great emphasis on accuracy, meaning that participants were clearly loss-aversive. Furthermore, $\hat{\lambda}$ values were slightly overestimated, $\hat{D}$ was estimated to be close to 1s instead of the 0.5s used, and the $\hat{\alpha}$ values were very close to the experimental ones. Generally, the model appears to capture the data without needing to use highly different experimental values. However, RA provided a similarly good model, with also reasonable parameters (Table 7.7), and a fitting that is almost as good as the fitting for RR_m (see Figure 7.8 for a comparison). In Figure 7.10 it is possible to see how RA provided a better fit for $\alpha = 0.25$, RR_m a better one for $\alpha = 0.75$.

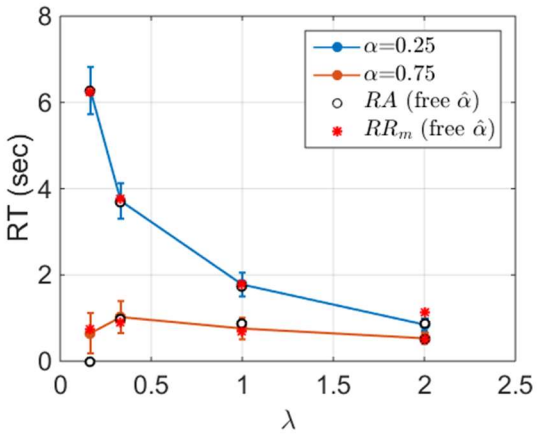


Figure 7.11

Median of median of responses, with the two best-fitting models (white circles and red asterisks).

We also tried a different model, by varying $\hat{q}$ across α conditions, as it is possible that participants adjusted their subjective weighting parameter according to the experimental condition. In particular, according to the "fourfold pattern of risk attitudes" in prospect theory (Tversky & Kahnemann, 1992), when they were in a "high reward probability" mindset (that is, when $\alpha = 0.75$), they may have had a more loss-aversive behaviour (corresponding in setting a high $\hat{q}$). We therefore used a version of RA, RR_m and BR with free $\hat{q}$ across α conditions, and we compared it with the fixed $\hat{q}$ estimated before. Note that RR does not have the q parameter. The results are shown in Table 7.9. In all these models, the parameter $\hat{\alpha}$ is left free to vary across α conditions.

Table 7.9 Results for the decision rules fitting (free or fixed $\hat{q}$)

	BIC	df
$\boldsymbol{RR_m\ (fixed\ \hat{q})}$	**-0.55746**	**8**
$\boldsymbol{RR_m\ (free\ \hat{q})}$	0.570953	9
$\boldsymbol{RA\ (fixed\ \hat{q})}$	-0.12927	8
$\boldsymbol{RA\ (free\ \hat{q})}$	0.205938	9
$\boldsymbol{BR\ (fixed\ \hat{q})}$	1.205431	7
$\boldsymbol{BR\ (free\ \hat{q})}$	1.334914	8
$\boldsymbol{RR}$	1.84737	7

As expected, allowing $\hat{q}$ to vary across α conditions provided a better WLS value. For each decision rule, the $\hat{q}$ pattern across the two conditions was clear: $\hat{q}$ was remarkably higher than 1 with $\alpha = 0.25$ condition, and slightly lower than 1 with $\alpha = 0.75$ condition, meaning that participants were giving more weight to the reward than to the punishment in the latter case. However, RR_m with fixed $\hat{q}$ again provided the best mWLS value, as the estimated $\hat{q}$ where too different from the experimental value of 1 (for RR_m, these values were 4.13 and 0.79 for $\alpha = 0.25$ and $\alpha = 0.75$ respectively). How relevant was this difference clearly depends on the penalty we set for deviating from real experimental

parameters, and is therefore open to interpretation. Again, RA with free q provided a similar good fit, whereas the other two functions provided an extremely poor fit.

To summarise, varying α had a large effect on the response, and a model with fixed $\hat{\alpha}$ cannot account for the observation. The best model appeared to be RR_m, and the estimated $\hat{q}$ value was much higher than the experimental one. It is unclear whether participant used different $\hat{q}$ for the 2 α conditions, as both versions provided a good account for the data. If this is the case, participants generally gave more weight to the punishment during the $\alpha = 0.25$ condition, and more weight to the reward during the $\alpha = 0.75$ condition, which is opposite of what expected from the Tversky & Kahnemann (1992) fourfold pattern of risk attitudes.

7.8.3 Analysis of performance groups

By dividing the dataset into three groups based on performance (i.e. point earned, see **Section 3.1**) we noticed a clear pattern (see Figure 7.11). Participants in the best performance group were faster, allowing them to obtain a higher reward with the $\alpha = 0.75$ condition, where the optimum was $RT^* = 0$. It appeared that participants in the other group were not just slower, but they may actually have different, subjective parameter estimations. In fact, if they were trying to reach $RT^* = 0$ without being able to do so for motor limitation, their response in the $\alpha = 0.75$ condition would always be equal to their fastest response for that condition. Instead, their response followed a precise pattern, (a monotonic increasing one for the poor performance condition and a non-monotonic one for the medium performance group). This may be due to a more loss-aversive behaviour, which could correspond to a high $\hat{q}$ value, as we will confirm when fitting the dataset.

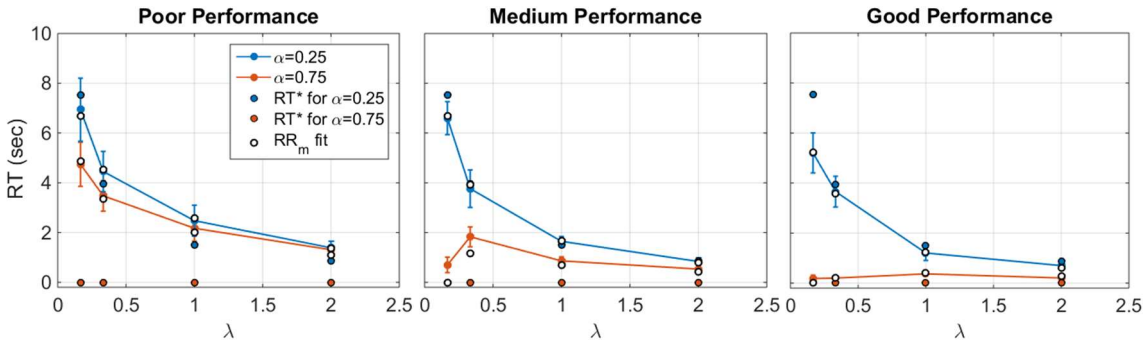


Figure 7.12

Median of median RT for each performance group (circles connected with lines), with optimum responses (circles with black edges) and best fitting model (white circles). For each performance group, the best fitting model resulted in being RR_m.

Interestingly, participants in the best performance group were faster even when this was not optimal. To measure that, we take the Euclidian Distance $\sqrt{\sum_{i,j}\left(mdRT_{i,j} - RT^*_{i,j}\right)^2}$ where the index i corresponds to each λ, and j to each α condition. The results are shown in Table 7.10. The Good Performance group, even though was much closer to the RT^* for the $\alpha = 0.75$ condition (as clearly seen from the figure), was more far away from the RT^* of the $\alpha = 0.25$ condition than the other two groups.

Table 7.10 Euclidian distance from optimum response

	Poor Perf.	**Medium Perf.**	**Good Perf.**
$\boldsymbol{\alpha = 0.25}$	1.354	0.9587	2.365
$\boldsymbol{\alpha = 0.75}$	6.402	2.2047	0.502

Fitting the decision strategies to performance groups provided some insight. We fitted RR, RR_m, RA and BR, but we excluded RR and BR as their fitting was always much poorer than the fitness for the other two. The preferred model (both in terms of BIC, mWLS, and WLS) was RR_m for all performance group, even though RA performed approximately the same for the poor and the average performance group (Table 7.11). There was also a slight decrease of the $\hat{q}$ in groups with higher performance, for RR_m and RA with the first two performance groups (Table 7.12). This means that participants in the higher performance group may have given more weight to speed compared to accuracy (for the RA

decision rule) or they may have decreased their subjective level of punishment/reward ratio (for RR_m), which would explain the totally different strategy they used (see top Figure 7.11).

Table 7.11 BIC value for each performance group

	BIC		
	Poor Perf.	**Medium Perf.**	**Good Perf.**
$\boldsymbol{RR_m}$	0.12144	0.677505	0.072034
$\boldsymbol{RA}$	0.12447	0.815835	0.608339

Table 7.12 Estimating $\hat{q}$ values for each performance group

	$\hat{q}$ values		
	Poor P.	**Medium P.**	**Good P.**
$\boldsymbol{RR_m}$	2.409	2.205	1.812
$\boldsymbol{RA}$	1.594	1.428	1.859

We conclude from this analysis that the difference between participants in different performance rules may reflect their subjective weighting parameters and, possibly, the strategy used. However, RA is a peculiar strategy because, even though does not maximize the number of earned points, is it a weighted sum of RR and accuracy (Bogacz et al, 2010). Unfortunately, the dataset for individual participants was not large enough to allow individual fitting of the data.

7.8.4 Distribution Analysis

The Vincentized distributions are shown in Figure 7.12, with the respective mean, standard deviation and skewness on each panel (note the different scale of the axes). As expected, most of the distributions presented a right-skewed shape, similar to what obtained with Classic RT paradigm (but the skewness was generally lower than for Classic RT distribution). Both λ and α appeared to have a significant effect on the standard deviation (for λ: $F_{3,57} = 32.336, p \cong 0$, for α: $F_{1,19} = 5.872, p = 0.026$) and a significant interaction ($F_{3,57} = 5.899, p < 0.001$). The α condition appeared to significantly affect the skewness ($F_{1,19} = 16.479, p < 0.001$), but λ did not ($F_{3,57} = 0.749, p =$

0.528). The rate of response was not normally distributed (by standardising and aggregating results from individual participants, both with and without using the STD-IQR method), confirming the results of Experiment 6.1 and 7.1.

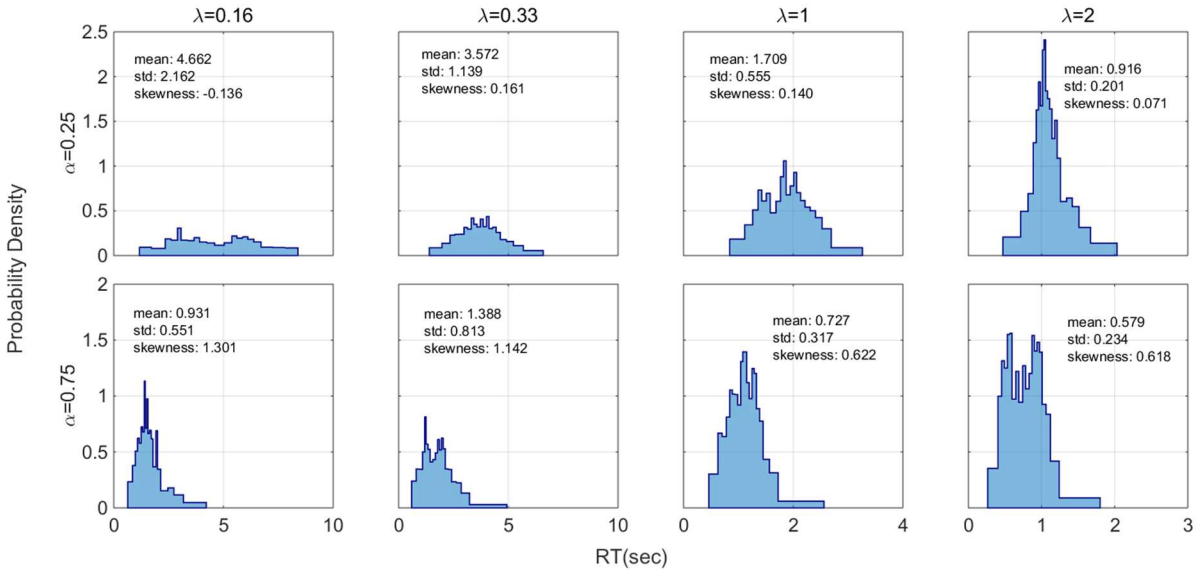


Figure 7.13

Vincentized distributions of responses for each condition. Note the different scales on the horizontal axis. As we found in the previous experiment, the standard deviation consistently decreases with increasing λ, apart for the condition with the lowest λ and $\alpha = 0.25$.

The accuracy of response was analysed by standardising and collapsing across all participants, both with the complete dataset and by excluding the most severely truncated distributions (see Methods in this Chapter, **Section 7.3**). In this case, the effect of modification appeared to be very mild (apart from two distributions in the $\alpha = 0.75$ condition). Most of the distributions were approximately normal as observed in the original dataset as well. The Kolmogorov-Smirnov test for normality did not reject the hypothesis of normality distribution for any distribution at significance level $\alpha = 0.01$ (Figure 7.14)

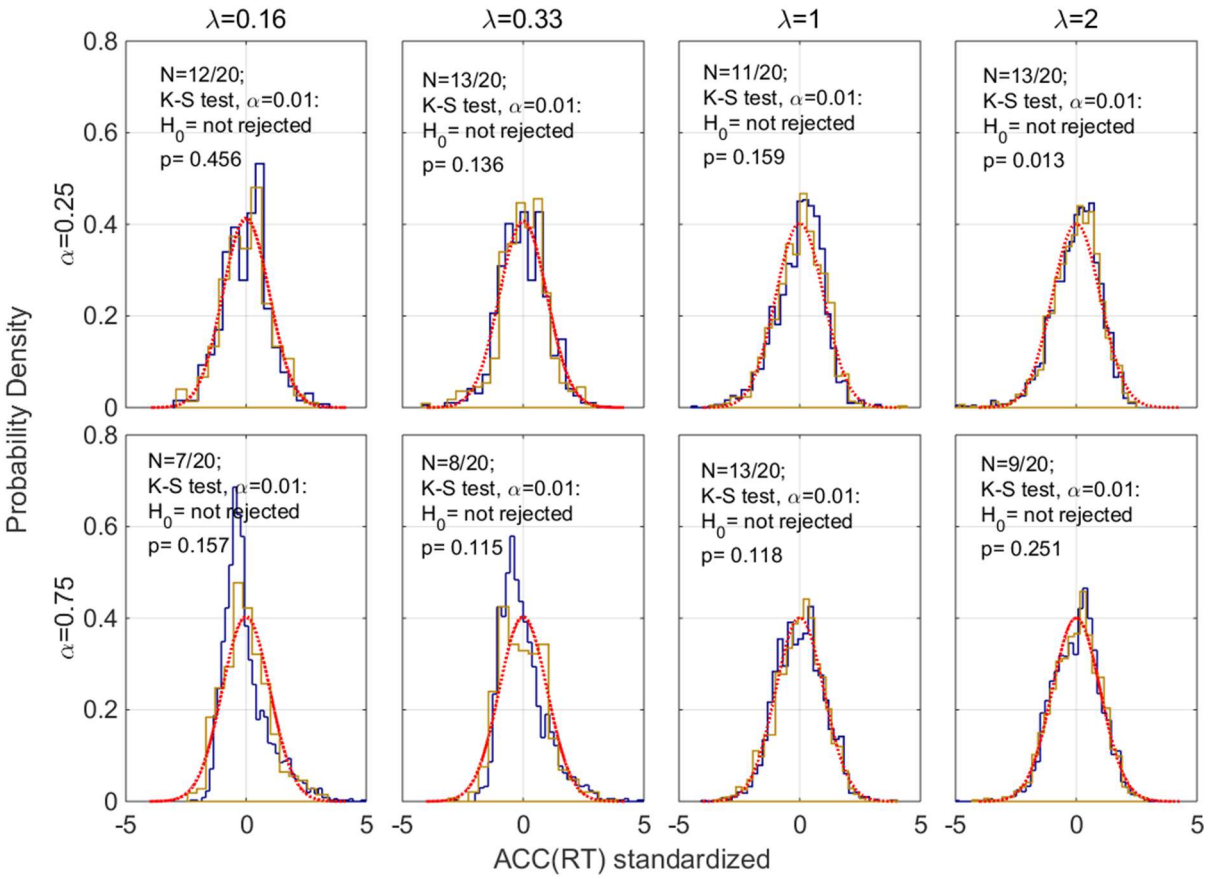


Figure 7.14

Distributions of $ACC(RT)$ responses. The blue lines correspond to the original dataset, the yellow lines to the dataset after the severely truncated individual distributions were eliminated. The number of eliminated datasets is shown in each panel (N). The result of the Kolmogorov-Smirnov test for normality is also shown in each panel, with the red-dashed lines corresponding to a Standard Normal Distribution.

This is consistent with the previous experiment and suggests, as we discussed previously, that at least in this paradigm participants were basing their response on the $ACC(RT)$, and it is tempting to speculate that something similar may happen with Classic RT as well. This will be discussed more deeply in the general discussion of this chapter (**Section 7.10**).

7.9 Discussion

In this study, we examined some of the prediction generated by coupling the exponential $ACC(RT)$ with the decision strategies that participants may use to maximize their subjective reward. Whereas in the first experiment of this Chapter we focused on the effect of different delay across trials on RT and accuracy (have already investigated in previous Chapters), in this section we tackled a new problem by

investigating the effect of changing the α parameter, which corresponds to varying the starting point at the beginning of the trial. This EXACT Paradigm proved to be a powerful resource for this analysis, as it allowed to vary α from the whole range 0-1, without affecting other parameters. The participants were clearly responsive to different α, producing RT^* curves that appeared to be in line with our predictions. In particular, low and high value of α produced a Piéron's and non-Piéron's shape, respectively. We also observed the Piéron's shape in Experiment 7.1 and this study confirmed that Piéron's law was the result of the combination of a speed accuracy function (which establishes the shape of $ACC(RT)$) and a maximization strategy (which establish at what point the accumulator of information must stop and the decision can be made). This contrasts with the idea that Piéron's law is a side-effect of the perceptual mechanism, as claimed by Stafford & Gurney (2004) within the context of DDM (see also Donkin & Van Maanen, 2014). The non-Piéron's shape, predicted by the model but not observed in the previous literature (neither in the context of Piéron's Law experiment or in the context of studies with different alternatives), shows us that participant can approximate complex decision strategies in order to maximize their reward.

By fitting the results to decision strategies, we had a better insight into participants decisions' mechanism. The optimum decision strategy was found to be RR_m, similar to previous study. For different performance group, the values obtained with RA were similar to RR_m. The $\hat{q}$ value was consistent across different participant's group and RA and RR_m strategies: it was generally higher than experimental value of 1 used and it was lower (but still higher than 1) in better performance groups. The relationship between performance and estimated $\hat{q}$ value considering the current literature will be more deeply investigated in the next experiment.

7.10 General Conclusions

In this Chapter, we explored several features of the new version of the EXACT Paradigm. Compared to the investigation in design in the previous Chapter, here we employed a study in which participants were completely aware of their accuracy in time, and could make choice accordingly without having to estimate it. We obtained a clear improvement in the quality of the dataset compared

to the previous set of experiments: participants' median responses were less spread out, there was much less response shifting and there was no relevant trend along conditions. Furthermore, by using several values of λ we could obtain several RT^* curves that we could fit with the decision strategies.

The main point of this Chapter was to show that optimality in RT tasks can indeed be measured without having to assume a particular perceptual process, by using an experimental setup that resembles a Classic RT tasks. The EXACT Paradigm could be used to investigate several other aspects of behaviour: the response to sound feedback (by following participants reports, this should influence how loss-aversive they behave, and possibly have an effect on the q parameter), the response to different speed-accuracy function shapes, the behaviour when $ACC(RT)$ fluctuates from a trial to another, instead than being equal across the whole experimental condition, as we assumed here. The effect of feedback on participants' responses will be thoroughly investigated in the next Chapter, by also providing a comparison with a Classic RT task.

In terms of optimality, we have again confirmed that participants were not basing their response on the delay across trials. We suggested in **Section 7.5** different interpretations for this. We stress again the importance of fitting data quantitatively in order to understand which decision strategy appears to be more likely used by participants: in fact, judging solely by the lack of effect of the delay D, we would have concluded that participants were using the BR strategy, as it is the only one among the four considered which is not related to trial time. However, BR fitting appeared unequivocally poor, whereas the RR_m model appear to be the preferred model.

This is not completely consistent with the results of Bogacz et al., (2010). In his study the participants were divided into 3 performance groups (Bottom 10%, Middle 60% and Top 30%), and the computed likelihood suggested that RA was the most successful model in accounting for the data (this was particularly strong for the bottom 10% of the participant, where RA was 43 times more likely of RR_m in generating the data). Both RA and RR_m where much more likely than RR. In our first experiment RR_m was always the preferred model, but the fit for RA was also good. Unfortunately, this

is the sole study in which different decision strategies were compared (for example, Simen et al., 2009, used only RR; Balci et al., 2011, compared RR and RR_m, but not RA), and therefore it is difficult to draw a definitive conclusion based on these results.

As we discussed previously, the RR_m is a particularly interesting model, as in most experimental conditions it can be considered the rule that maximizes the participants' subjective value. This rule seems to apply even when an explicit punishment is not provided, by assuming that participants have some internal, subjective punishment that makes them want to avoid incorrect responses. However, RR_m is the optimal rule only in task with fixed condition duration. When the duration is variable and the number of trials is fixed, the optimum response according to RR_m is to wait for an infinite amount of time, which is absurd. It is possible that, with variable condition duration, participants were using a different strategy (for example, adding a cost to waiting, similar to what assumed in Drugowitsch et al., 2012), or assuming some arbitrary condition length. This remains untested.

Finally, in this Chapter, we suggest a novel way of analysing the accuracy distribution. We observe how accuracy appeared to be approximately Normally distributed, which suggest that participants were actually be basing their response on some particular accuracy values, fluctuating around it. In this interpretation, the RT distribution shape is a by-product of normality distributed accuracy value. In the EXACT task accuracy is defined between α and 1, and in a Classic RT with 2 choices the lower boundary is 0.5. This implies that, depending on the mean accuracy position that participant choose, the resulting distribution may be more or less truncated. The resulting RT shape can be found to be equal to:

$$RT(t; RT^*, \lambda, \alpha) \qquad Eq.7.1$$

$$= \frac{\phi(1 + (\alpha - 1)\exp(-\lambda t), ACC(RT^*), \sigma)(1-\alpha)\lambda \exp(-\lambda t)}{\boldsymbol{\Phi}(1, ACC(RT^*), \sigma) - \boldsymbol{\Phi}(\alpha, ACC(RT^*), \sigma)}$$

Where ϕ is the normal probability function and $\boldsymbol{\Phi}$ (bold) is the cumulative normal probability function. Note that the distribution of RT is function of the optimum response RT^*, and of the corresponding accuracy, $ACC(RT^*)$. Thus, it depends on the decision rule used.

For the exponential $ACC(RT)$, this correctly results in a right skewed distribution. Also, the RT distribution depends on the $ACC(RT)$ parameters λ and α, as shown in Figure 7.14. For increasing λ or α the RT distribution become less spread and the mean decreases (fast responses), which is exactly what we observed for our distributions (see Figure 7.6 and 7.12). At least for λ this is what we expect intuitively also for Classic RT, as higher λ corresponds to higher stimulus intensity. For α, is difficult to compare this with observed distributions, as we are not aware of any paper in which a comparison between distribution from task with different alternatives is reported.

This Accuracy-Hypothesis can be tested empirically, by calculating the accuracy distribution for each participant fitting a truncated normal on it, and then obtaining the resulting RT pdf through Equation 7.1. Unfortunately, the study was not designed to test this prediction, and the sample size of individual participants is too small to allow such procedure. Our method of aggregating participants is useful when checking general distribution shape, but distribution truncation, and the expected variability in mean across participants, did not allow us to use this method for fitting RT empirical distributions (recall that we standardised accuracy distributions before aggregating them together). It would be also extremely interesting to test this hypothesis in Classic RT paradigm. However, as we have mentioned in the introduction, this is extremely troublesome, as Classic RT tasks do not allow a direct measurement of $ACC(RT)$.

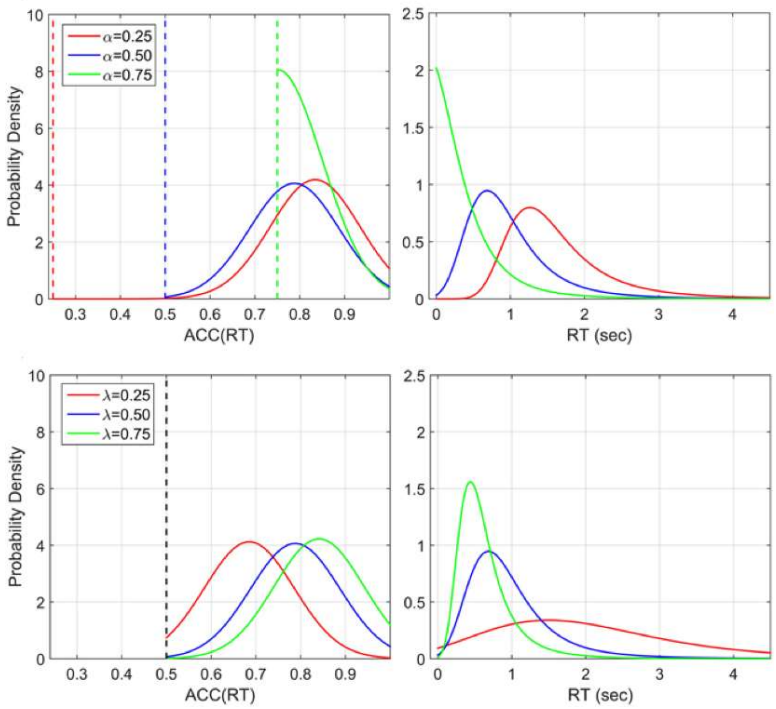


Figure 7.15

Possible explanation for the right-skewed distributions in the time domain, which would be generated by the Normal Distribution in the accuracy domain. The upper left panel show three distributions with the same λ (with $\lambda = 0.5$) and three different values of α, whereas the bottom left panel showing three distributions with the same level of $\alpha = 0.5$ and three levels of λ. The mean of the Normal Distribution is calculated to be $ACC(RT^*)$, with the RR_m decision rule. Note that, on the accuracy domain, α correspond to the lower truncation point of the accuracy distribution. Different values of λ and α correspond to different RT distributions.

Chapter 8

Payoffs Effect for the EXACT Paradigm and Classic RT Task

8.1 Introduction

In the previous Chapter, we explored the capability of the EXACT Paradigm. The main aim of this Chapter is to compare it with a Classic RT task by investigating the effect of payoffs on RT and accuracy. Both experiments were organised similarly to the previous EXACT Paradigm task: we varied the difficulty of the trial (by changing the λ for the EXACT Paradigm and the stimulus contrast for the Classic RT task) and the weighting parameter q. In this way we were able to obtain different RT curves.

As we already discussed, even though the parameter q can be generally seen as an "emphasis on accuracy" parameter, recall from **Section 2.6** that this parameter have slightly different interpretation for different decision rules: it corresponds to the ratio between the cost of error over the cost of waiting for BR; the subjective utility for an incorrect response over the subjective utility over a correct response for RR_m (punishment/reward ratio); the cost of error (scaled by total trial time) and the weight given to reward rate for RA. In both the EXACT Paradigm and Classic RT task, we decided to follow the interpretation of the RR_m decision rule, and modify the punishment to reward ratio by changing the number of points the participants gain or lose when a correct/incorrect response was made. The first reason behind that is that, in previous studies, RR_m has been shown to be the strategy that could fit experimental observation better, even across different performance groups. Secondly, the interpretation of the parameter q within the RR_m is much clearer than within the RA or BR, as it is directly connected with the reward and punishment used, which is easy to manipulate. Nevertheless we will not exclude these other two models from our analysis in fitting the experimental data. In fact, we believe that manipulating the punishment/reward ratio may still have an indirect effect on the q parameter for the

other decision rules, as it is affecting the emphasis on accuracy. As we have seen in **Section 2.8**, the general effect of increasing q in terms of DDM is to increase the optimal threshold, as higher threshold increases accuracy.

8.1.1 Payoff Matrices

There are several ways in which payoffs can be used to affect behaviour. Participants may be differentially rewarded according to their reaction times, by giving trial by trial feedback, or different payoffs may be used for different stimulus-response combinations. (e.g. Luce, 1986; Snodgrass et al., 1967; Saslow, 1968, 1972). The way payoffs are organised in a task can be summarised through payoff matrices, like those used in Table 8.1, where A or B represents the correct alternative (in a 2 choices task), and a or b represents the participant's choice. For example, the cell corresponding to A/a indicates the feedback given for a correct response to the alternative A; the cell B/a indicates the feedback given for an incorrect response, with B being the correct alternative. There are two main types of payoff matrices. The first type regards matrices used when the aim is to affect the response bias. This is the case when the rewards and punishment are given such that participants will favour one alternative instead of the other, as shown in Table 8.1 (Type 1). For example, by setting the sum of the feedback for A/a and B/a higher than the sum for A/b and B/b we can bias the participants towards the response "a". We define matrices of Type 2 those used to affect the speed-accuracy trade-off, by giving unbalanced feedback for correct and incorrect responses: in Table 8.1 (Type 2), we usually set Y below 0, so that an incorrect response corresponds to punishments (however, both positive and negative values could be used).

Table 8.1 Three types of payoff matrices

Type 1	**A**	**B**
a	X	Y
b	Z	W

Type 2	**A**	**B**
a	X	Y
b	Y	X

EXACT	$\boldsymbol{U > ACC(RT)}$	$\boldsymbol{U < ACC(RT)}$
R	X	Y

The third matrix in Table 8.1 corresponds to a possible representation of the matrix for the EXACT Paradigm. In this case, U is a random number drawn from a Uniform Distribution. If this number is higher than $ACC(RT)$ for a participant response (R), a feedback X is given, otherwise a feedback Y is given. By noticing that $U > ACC(RT)$ corresponds to a correct response, and $U < ACC(RT)$ to an incorrect one, we can see that this matrix is equivalent to the Type 2 matrix. However, the EXACT Paradigm cannot be used to test prediction for "biased" responses (Type 1). In order to have our result consistent with the EXACT Paradigm study presented earlier, we are going to use the type of payoff matrix represented in Table 2.

Most of the studies focused on the response bias, using, therefore, the first type of payoff matrix (Diederich & Busemeyer 2003; Diederich, 2008; Liston & Stone, 2008, Simen et al., 2009 [exp.3], Krajbich, Armel & Rangel, 2010; Milosavljevic et al., 2010). Within this type of payoff matrix, when applied to sequential sampling models, three hypotheses have been formulated (Diederich & Busemeyer; 2006): the bound-change hypothesis assumed that participants change their threshold bias (the distance between the starting point and the threshold of the sequential sampling model) by setting the starting point of the process closer to the alternative with higher reward. The Drift-Rate-Change hypothesis assumed that the drift rate for each alternative depends on the sum of evidence accumulated and value for that hypothesis. Finally, the two-stage-process (Diederich, 2008) suggests that the task may be seen as a multi-attribute decision problem, with two stages: first deciding the payoff and then the stimulus evidence. It turned out that this last model provided the best account for the data. However, in a similar experiment, Leite (2012) has shown how both a two-stage diffusion model with a shift in drift rate and a single-stage diffusion model with shifts in starting point could account for the data, and the two models were so similar that it was not possible to choose between them.

After the theoretical presentation of Bogacz et al. (2006), researchers have focussed their attention on optimum performance with biased alternatives (two alternatives, see Type 1 matrix in Table 8.1). Within the context of pure DDM, this allowed much more precise predictions: estimated starting point of the process was predicted to be closer to the more rewarded alternatives, the average RT should be

shorter and accuracy lower for that response, and large reward asymmetries with sufficiently short delay across trials should generate a non-integrative response (that is, a response corresponding to the biased rewarded alternative, as fast as possible, and with accuracy at chance level). Simen et al. (2009) tested these assumptions, by using a 2AFC random-dot motion task with 3:1 reward/punishment ratio, finding that the predictions were satisfied for most of the participants.

Even though the studies based on matrices of Type 1 have been investigated across different frameworks (e.g., for a signal detection approach, see Maddox & Bohil, 1998), a systematic investigation for matrices of Type 2 does not exist. This is not to say that these types of payoff are not used: in several cases, we found that they have been implemented in the design as a way to discourage people from committing mistakes and/or anticipations, by using monetary punishment (e.g. Briggs & Blaha, 1969) or some other types of negative feedback. (Jones et al., 2013). This procedure is so common that, coupled with the (according to what we observed) inherently participant's risk-avoidance, the subjective value of the punishment appears to be almost always much higher than the value for the reward.

In one of the few studies where a matrix of Type 2 was used (Blank et al., 2013), by using a perceptual decision task and three different punishment levels, it has been found that higher punishment appeared to increase the accuracy of response, but strangely not RT, which is difficult to explain in terms of DDM. In fact, increase in accuracy can be due to higher drift rate or higher threshold, but this would correspond to slower or faster RT, respectively. It also found that there was a different for the 10^{th}, 50^{th} and 90^{th} percentile of the behavioural effects defined as "punishment – no-punishment condition" in accuracy, finding a negative value for the low percentile and a positive for the high percentile, suggesting that, in some participants, punishment impaired performance, whereas in others, it improved it.

Instead of using different payoff matrix, Balci et al. (2011) analysed how the q parameters changes along with training. They found that it decreases from 0.8 to 0.2, corresponding to an increase in subjective reward and a decrease in the normalized threshold for the Pure DDM. Bogacz et al. (2010)

found that the estimated value of the$q$ parameter was connected with the performance of the participants, with lower q values corresponding to higher performance (more details about these experiments will be given in the second study of this Chapter).

*8.1.2 Relationship between q Parameter and λ and the Corresponding RT**

The relationship between the q parameter across different λ on the corresponding RT^* is similar to that observed with α, as these two parameters are related. This relationship has already been discussed informally in **Section 2.10**, and in more detail in **Section 6.1.5** (see also Figure 2.5). Theoretically, with $\alpha = 0.5$ participants' response should assume a Piéron's shape whenever punishment is higher than reward, and a non-Piéron's otherwise. In particular, recall that for RR_m we obtain a non-Piéron' shape (for different values of λ) whenever $\alpha > \frac{q}{q+1}$. With $\alpha = 0.5$, this is true whenever $q < 1$ (reward higher than punishment). In these cases, Piéron's Law fits the data well ($R^2 \cong 0.999$). We found that the best Piéron's variant for these RT^* curves was the one with free k and β with both k and β increasing with increasing q, and a $\gamma = 0.12$. Is interesting to note that, even though mathematically the function goes asymptotically to 0 with increasing λ, this is not captured by the Piéron's fitting when value of λ up to 10 are used (when we fitted a simulation with value λ as high as 100, the correct $\gamma = 0$ was estimated), which may lead to assume that a similar mechanism may be going on in observed human RT. We obtained the same results with varying α, which resulted in a $\gamma = 0.083$ (**Section 7.2**).

When the subjective punishment/reward ratio is exactly equivalent to the amount of points given/taken, then is predicted that the participants will generate a non-Piéron's shape whenever the points given for a correct responses are higher than the point taken for an incorrect one ($q < 1$). However, as we discussed previously and observed in the previous study, it appears that participants are normally loss-avoiding, which corresponds in giving more value to the punishment and therefore having a subjective q value generally higher than 1. We also believe that this is one of the reasons the non-Piéron's shape has not been observed in Piéron's-like experiment. To reverse this loss-avoiding strategy, we need rewards that are much higher than the corresponding punishment, as the exact mapping between points and subjective q value is not known.

8.1.3 Organization of the Chapter

One of the main aims of the final two experiments was to investigate the effect of manipulating the punishment/reward ratio. In particular, we wanted to test the prediction about Piéron's and non-Piéron's shape with EXACT and Classic RT task, and to compare the results across both experimental designs. In a Classic RT Paradigm, apart from testing the Piéron's and non-Piéron's law prediction, we also tested for different decision strategy by assuming the DDM as the perceptual model used. We made several assumptions about the estimated threshold, reaction time, error rate, etc. As this is not relevant for the EXACT Paradigm, we will discuss these predictions both intuitively and mathematically in the Introduction and Method section for the next experiment (**Section 8.6** onwards).

EXPERIMENT 8.1

The first experiment we investigated the effect of varying the punishment/reward ratio by using the EXACT Paradigm. The design was similar to the one used in the previous Chapter.

8.2 Methods

We recruited 31 participants (14 males, 17 females) to take part in this experiment. A gauge was shown on the screen, along with the points earned or lost for the last trial (in green if earned, in red if lost). The gauge increased in time according to Equation 6.1, which established the speed-accuracy trade off $ACC(RT)$. Participants were asked to press CTRL whenever they wanted to, whereupon a reward was given with probability $ACC(RT)$. When a reward was not given, a punishment was given instead. After each participant's response, we set a pause of 0.5 seconds ($D = 0.5s$) and then the trial started again (the gauge was reset). See Figure 7.1. We kept $\alpha = 0.5$, meaning that the gauge started half full at the beginning of each trial. We used 4 different gauge speeds ($\lambda = 0.16, 0.33, 1, 2$, higher λ corresponding to faster increase of the gauge) and 3 different q conditions. We indicated the q conditions by referring to the punishment/reward used, thus $q = \frac{10}{10}, \frac{10}{1}, \frac{1}{10}$. For each q condition, we ran all the λ level (blocked, in randomized order) before proceeding to the next q condition. However,

the first q condition was always set up to be $q = \frac{10}{10}$, in order to get participants accustomed with the paradigm on a balanced level of punishment/reward, whereas the other two q conditions were randomized. Before each q condition, participants performed a short training with the corresponding q value and $\lambda = 1$. Before each session an example of the progressing gauge was shown. Each session lasted 2 minutes. Participants started each session with 100 points. Each participant was paid £4 for their participation plus an amount of money depending on their performance (25 points=1 pence).

8.2.1 Predictions in terms of Optimum Response

Note that the optimum response time (by using RR_m) calculated using the experimental q does not result in a smooth shift between Piéron's and non-Piéron's shape as in Figure 6.2 (reproduced here as Figure 8.1).

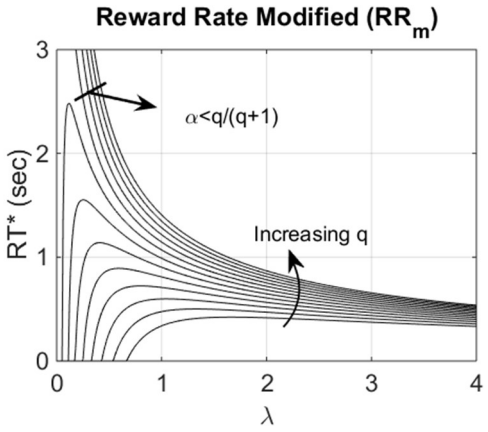


Figure 8.1

Several RT^* curves, generated by increasing the value of q (the emphasis on accuracy parameter). Note that some of these curves go to infinity with λ going at zero (indicated by the black straight line), depending on the combination of q and α parameters.

Instead, the optimum RT^* generates a Piéron's shape for $q = \frac{10}{1}$ and $= \frac{10}{10}$. For $q = \frac{1}{10}$ the RT^* curve is equal to 0 for all λ but the last one, in which is $RT^* \cong 0.1s$, for $q = \frac{1}{10}$ (very high reward), as shown in Figure 8.2. A clear non-Piéron's shape like the one observed in the simulation can be obtained by using a less rewarding condition, for example $q = \frac{1}{5}$. The reason we did not chose this value was that, based on what found on previous experiment, and after having ran several pilot participants on this

study, we concluded that participants gave generally a higher weight on the punishment. Therefore, using $q = \frac{1}{5}$ may not be rewarding enough to generate a non-Piéron's shape (as this q condition may correspond to a higher subjective punishment/reward ratio, even higher than $q = 1$ if participants were not very loss-aversive). We therefore used an even lower q that it would allow us to obtain non-Piéron's shape. As we will see, things are more complicated as the shape will depend on the performance of the participant.

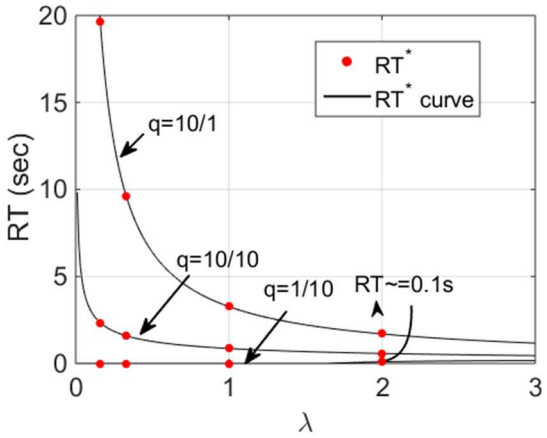


Figure 8.2

The three different RT^* curves (in black) obtained with the three levels of punishment/reward ratio used in the experiment. The red dots indicate the RT^* values for the λ used in the experiment ($\lambda = 0.16, 0.33, 1, 2$). Note that the optimum time to respond for the $q = \frac{1}{10}$ condition is equal to 0 for the first three λ conditions.

8.3 Results

We analysed the total score for each participant/condition and compared it to the theoretical average score assuming that participants were trying to earn as many points as possible (Figure 8.3). In the $q = 10/10$ and $q = 10/1$ conditions participants appeared to earn approximately the same amount of point as predicted by the theoretical average score. For $q = 1/10$, instead, they consistently earn less point than the average optimum, due to the fact that they are slower than the theoretical optimum response (which is $RT^* = 0$ in this case). As in the previous study, it is possible that participants did not reach the optimum score as they were not able to reach such short RT, or it was deemed to be inconvenient (producing so fast RT may have a cost, as we observed in **Section 7.4)**. Another possibility is that participants were using a very different estimate for q. As before, note that an increase in participant's

score does not necessarily corresponds to higher RT, as this increase may be simply due to higher λ which resulted in more trials.

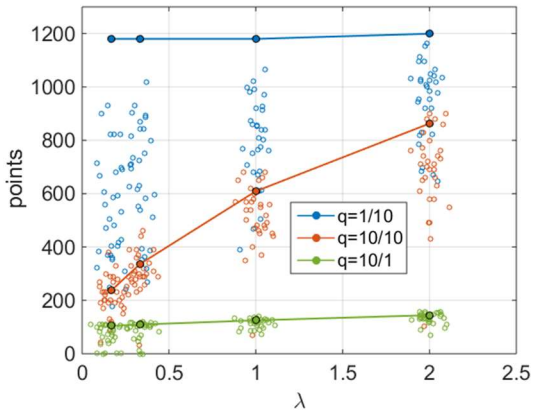


Figure 8.3

Empty circles indicate the total score for the whole experiment for each individual participant, whereas circles with black edges (connected with a line for each q condition) correspond to the expected points obtained by using the optimum response. For the $q = \frac{10}{1}$ and $q = \frac{10}{10}$ conditions participants are mostly close to expected scores, whereas they are clearly sub-optimal for the $q = \frac{1}{10}$ condition.

8.3.1 Response Signal Analysis, Trends, and Feedback Adjustment

The analysis of the number and magnitude of shifting of responses in each condition by using the STARS method (**Section 3.9**) did not present any relevant surprise: the number of shift was fairly stable at around 2 per condition with a slight but not significant increase for higher λ conditions, and a slight and non significant decrease for higher q conditions; the average magnitude of the shift equalled 0.821, without any significant trend for different conditions.

By aggregating response signal across participants, we noticed the presence of several trends across q conditions. In particular, there was a fairly consistent, decreasing trend for the $q = \frac{10}{1}$ condition and an increasing trend for the $q = \frac{1}{10}$ condition. The trend for the $q = \frac{10}{10}$ condition was, instead, less consistent across λ conditions (see Figure 8.4).

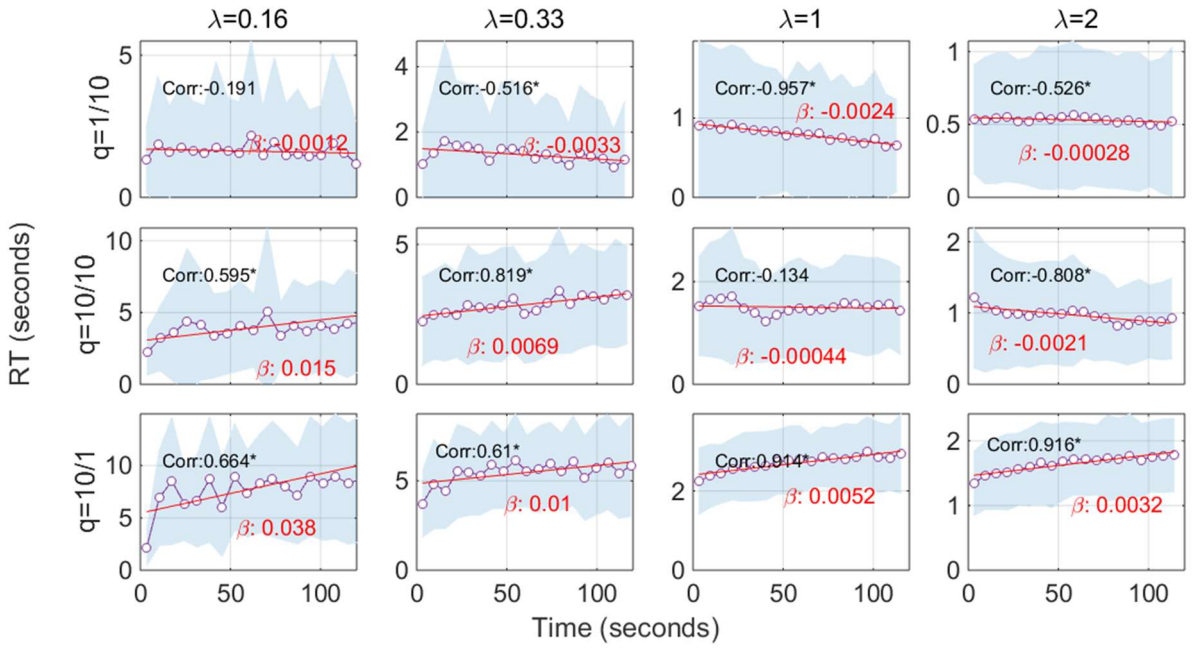


Figure 8.4

Aggregated response signal across participants for each condition. Significant correlations are indicated on each panel by an asterisk. A linear function is fitted on each panel (red line), with β indicating the slope of the line.

It appeared that participants were adjusting their response along each session, based on the q condition. In particular, they became slower in time with the high punishment conditions, and faster with the high reward conditions. Is it possible that participants started with some prior value of $\hat{q}$ which is close to $q = 10/10$. Participants may adjust their estimate $\hat{q}$ based on the feedback during the session. In spite of this, the analysis of the change in standard deviation across condition revealed again no correlation: participants were adjusting their response, but their response was not getting less sparse as time goes on. Of course, these analyses were limited by the very short length of each condition, and it is possible that a reduction in standard deviation may be observed with longer sessions.

The analysis of adjustment of responses after punishment or reward did not produce, again, any surprise, by confirming the results from the previous studies: adjustment in response after rewards was remarkably smaller than adjustment after punishment (change after rewards was almost 0, whereas average change after punishments was ~0.025s). Moreover, as predicted, adjustment (after punishment) were higher with higher punishment/reward ratio (average of 0.006s, 0.015s, 0.025s for the conditions $q = \frac{1}{10}, q = \frac{10}{10}$ and $q = 10/1$ respectively (the results were, however, non-significant).

8.3.2 Effect of q and λ Condition on RT and ACC(RT)

The results in terms of median of median RT (mdRT) and accuracy of response ($ACC(RT)$) are shown in Figure 8.5 for each q condition. The optimum response for each condition is also shown (circles with black edges). Notice that for $q = 1/10$ the optimum response was always $RT^* = 0$ (resulting in $ACC(RT) = 0.5$), except when $\lambda = 2$, in which case the optimum response was roughly $0.1s$. Both λ and q affected significantly the median RT (for λ: $F_{3,90} = 57.805,\ p < 0.0001$, for q: $F_{2,60} = 58.794,\ p < 0.0001$). We also founded a significant interaction between the two: ($F_{6,180} = 21.858, p < 0.0001$). The effect of λ and q condition was similarly significant on $ACC(RT)$ (for λ: $F_{3.90} = 74.066,\ p < 0.0001$, for q: $F_{2,60} = 79.469,\ p < 0.0001$). Conversely to Experiment 7.2, where we were similarly investigating Piéron's and non-Piéron's shape, the interaction was in this case not significant ($F_{6,180} = 1.535, p = 0.169$). Post-hoc pairwise comparisons by applying Bonferroni correction revealed that both mdRT and $ACC(RT)$ were significantly different for each λ and q condition ($p \cong 0$). Overall, participants appeared to be responsive to manipulating the punishment/reward ratio by changing both time of response and accuracy (positional response along the gauge). Compared to the optimum theoretical response, participants were faster with the high reward conditions ($q = 10/1$), and slower with the high and balanced punishment conditions ($q = 1/10$and $q = 10/10$ respectively). This suggests a complicated relationship between experimental and subjective punishment/reward ratio, which will be analysed in the model fitting section.

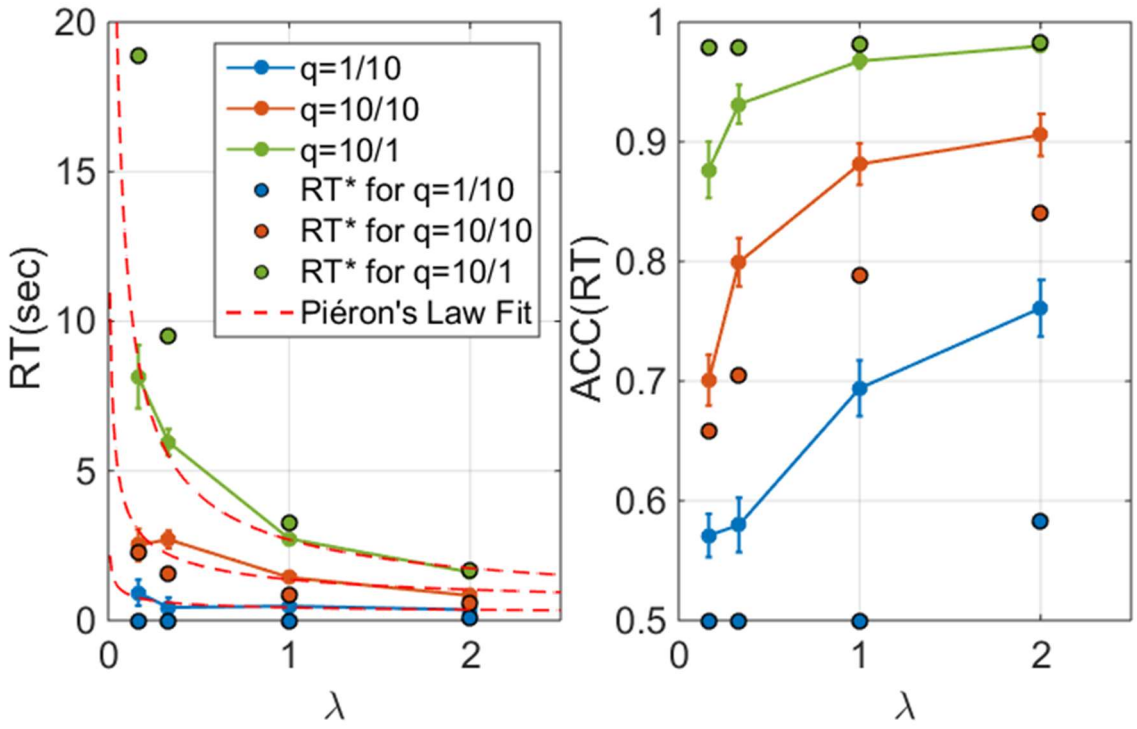


Figure 8.5

The circles connected with lines indicate the median of median RT for each condition (**left panel**) and corresponding $ACC(R$ (**right panel**). The circles with black edges corresponds to the optimum response. The red-dashed lines correspond to the Piéron's Law fitting, with a different Piéron's function for each q condition.

We fitted Piéron's Law by assuming that k, β, γ, any combination of those (or none of them) were free for each condition, thus comparing 8 different models. The model with lower AIC was the one with k and β free to vary across q conditions (Table 8.2). In particular, both k and β were found to be increasing across increasing q conditions (Table 8.3). This matches the predictions generated by our simulation (see **Section 8.1.2**). The fitting is shown as a dashed red line in Figure 8.5.

Table 8.2 Fitting Piéron's Law Models

Model	BIC	df
No free	1.286628	3
k free	-1.58119	5
β free	0.98539	5
γ free	0.8117	5
k and β free	-1.52056	7
k and γ free	-1.43845	7
β and γ free	-0.16903	7
k, β and γ free	-1.20063	9

Table 8.3 Parameters for the best model

Model 1	
BIC	-2.85043
ΣWLS	1.4358
$\boldsymbol{k}_{q=\frac{1}{10}}$	0.283
$\boldsymbol{k}_{q=\frac{10}{10}}$	1.224
$\boldsymbol{k}_{q=\frac{10}{1}}$	2.55
$\boldsymbol{\beta}_{q=\frac{1}{10}}$	0.427
$\boldsymbol{\beta}_{q=\frac{10}{10}}$	0.473
$\boldsymbol{\beta}_{q=\frac{10}{1}}$	0.674
$\boldsymbol{\gamma}$	0.148

The goodness of fit in terms of R^2 was equal to $0.702, 0.779$ and 0.98 for $q = \frac{1}{10}, \frac{10}{10}$ and $\frac{10}{1}$ respectively. As expected, the RT fits poorer and poorer with Piéron's Law as q increases, which is consistent with the idea that at high q the optimum decision strategy assumed a non-Piéron's shape. However, the results were not as clear as expected. In particular, we observed a slight decreasing trend with $q = \frac{10}{10}$, but not with $q = \frac{10}{1}$. Compare the empirical data with the simulation generate in Figure 8.1 and 8.2). We do not observe a clear decrease for the lower λ value, like the one observed when manipulated the starting point (α) in the previous Chapter. By looking at individual differences, we observe that some of the participants have this inflection points, and some do not.

8.3.3 Model Comparison for Aggregated Data

We fitted the BR, RA, RR_m and RR model to the aggregated data in Figure 8.5 by penalizing the fitting based on the distance between experimental parameters (λ, q, d and α) and subjective ones ($\hat{\lambda}, \hat{q}, \hat{d},$ and $\hat{\alpha}$). We used the same procedure as in the previous Chapter (see **Section 7.3**). We firstly compared, for each model, a version with free $\hat{q}$ or fixed $\hat{q}$ (RR does not have a q parameter, so it can only be compared with the other models, not with a different version of itself). Each variant with fixed $\hat{q}$ was remarkably worst (in terms of both BIC, sWLS and mWLS) than the variant with free q (Table

8.4). This was not surprisingly as the fixed $\hat{q}$ models consisted of just one RT curve which is compared with the three observed RT curves. As a further check, we also compared these models with a variant with fixed $\hat{q}$ and free D across q conditions: this gave some freedom to the RR model and allowed us to check whether the punishment/reward ratio really affected the q parameter and not some other parameters, like D. As expected, the fitting for these new variants were in all cases extremely poor, with BIC values higher than 6. Therefore, we can be reasonably confident that varying the punishment/reward ratio effectively affected the $\hat{q}$ parameter.

By comparing models with free $\hat{q}$, it is interesting that the best model appeared to be RA instead of RR_m as found in the previous Chapters. The fitting in terms of WLS for RR_m appears to fit the data reasonably well, as shown in Figure 8.6, but some estimated values were unreasonably different to the respective experimental values (see estimated values in Table 8.5). For example, $\widehat{D}$ wass estimated to be equal to 1.823s, much higher than the experimental value of $0.5s$. Compare this with the more reasonable estimate of $0.821s$ obtained with RA. However, both models captured the increasing in the $\hat{\lambda}$ parameter across increasing λ conditions, and the increase in $\hat{q}$ across increasing q conditions. For RR_m, the estimated $\hat{q}$ were always higher than 1, which implies that participants were weighting punishment more than accuracy. With RA, the interpretation of the $\hat{q}$ parameter is similar: $\hat{q}$ values lower than 1, obtained with $q = 1/10$ (high reward condition) and $q = 10/10$ (balanced condition) indicates a higher weight on reward rate, whereas values higher than 1, obtained with $q = \frac{10}{1}$ (the high punishment condition) indicates a higher weight on accuracy.

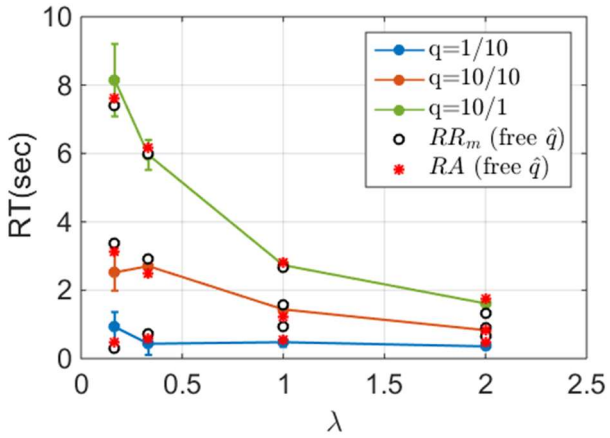


Figure 8.6

Median of median of responses, as shown in the previous figure, with the two best model shown with white circles and red asterisks.

Table 8.4 Results for the decision rules fitting

	BIC	df
$\boldsymbol{RR_m\ (free\ \hat{q})}$	0.637083	8
$\boldsymbol{RR_m\ (fixed\ \hat{q})}$	2.13685	7
$\boldsymbol{RA\ (free\ \hat{q})}$	**-0.50024**	8
$\boldsymbol{RA\ (fixed\ \hat{q})}$	2.141933	7
$\boldsymbol{BR\ (free\ \hat{q})}$	1.242078	7
$\boldsymbol{BR\ (fixed\ \hat{q})}$	2.03323	6
$\boldsymbol{RR\ (no\ \hat{q})}$	2.0944	7

Table 8.5 Best models' parameters

$RR_m(free\ \hat{q})$		$RA(free\ \hat{q})$	
$\hat{\lambda}_{\lambda=0.16}$	0.299	$\hat{\lambda}_{\lambda=0.16}$	0.703
$\hat{\lambda}_{\lambda=0.33}$	0.378	$\hat{\lambda}_{\lambda=0.33}$	0.843
$\hat{\lambda}_{\lambda=1}$	0.963	$\hat{\lambda}_{\lambda=1}$	1.661
$\hat{\lambda}_{\lambda=2}$	2.224	$\hat{\lambda}_{\lambda=2}$	2.585
$\hat{\alpha}$	0.698	$\hat{\alpha}$	0.459
$\hat{D}$	1.823	$\hat{D}$	0.821
$\hat{q}_{q=\frac{1}{10}}$	1.219	$\hat{q}_{q=\frac{1}{10}}$	0.199
$\hat{q}_{q=\frac{10}{10}}$	2.568	$\hat{q}_{q=\frac{10}{10}}$	0.749
$\hat{q}_{q=\frac{10}{1}}$	7.068	$\hat{q}_{q=\frac{10}{1}}$	5.685

8.3.4 Analysis of Performance Groups

By grouping participants based on their performance, we can shed some light onto these results (Figure 8.7). We firstly confirmed that, as expected, the good performance group was on average faster than the medium performance group, which was, in turn, faster than the poor performance group. However, the good performance group was slower than the other two groups in the condition with $q = \frac{10}{1}$ and lower λ, which is also a better strategy as the optimum point is, in that condition, extremely high (~19 seconds). Interestingly, the group with poorer performance presented clearly the non-monotonic shape for both the $q = \frac{10}{10}$ and $q = \frac{1}{10}$ conditions, and not for the $q = \frac{10}{1}$ condition, as expected. However, the same non-monotonic shape was not clearly present in the other two performance groups. From our dataset is not possible to establish if this is due to a difference in the strategy (or in the estimated parameters) used by the participants in a certain performance group, or it depended on a noisy dataset.

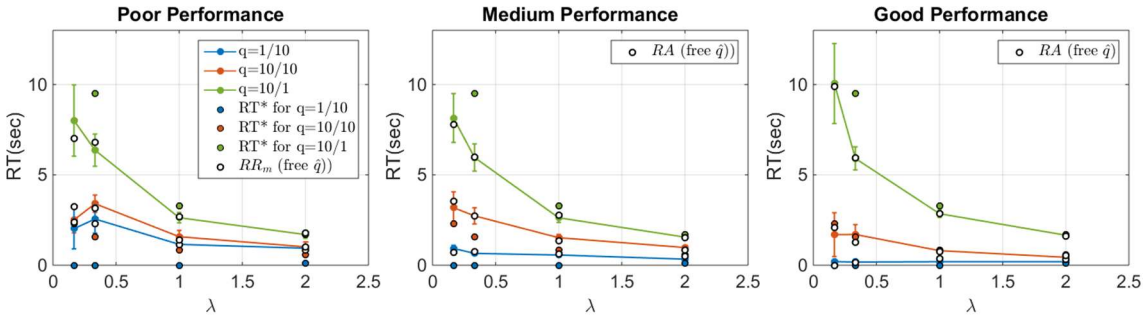


Figure 8.7

Median of median RT for different performance group (circles connected with lines), optimum response time obtained with RR_m (circles with black edges) and best fitting model (white circles). Note that the best fitting model is RR_m for the Poor Performance group, and RA for the other two groups.

By fitting different performance groups, we observed that the preferred model (in terms of modified WLS, WLS and BIC) was RR_m (with free $\hat{q}$) for the poor performance group, and RA (again with free $\hat{q}$) for the other groups. The BR and RR model performed very poorly and thus they are not reported (see Table 8.6) Even though the RR_m model did not provide the best fit, it was still reasonably good and the estimated parameters were close to the experimental ones. The $\hat{q}$ parameters for the RR_m in the poor performance group were estimated to be always higher than 1. The fitting for the BR model was consistently worse across the performance groups and, for both RA and RR_m models, the $\hat{q}$ values were always increasing for increasing q condition (as shown in Figure 8.8). Even though, for both RA and RR_m condition, better performance groups appeared to have estimated $\hat{q}$ closer to the experimental one, we could not detect any pattern along different performance group: that is, the group with worst performance did not appear to be on average more risk aversive than the other groups, but it just appears to have internal punishment/reward ratio which differs from the experimental punishment/rewards, which led to poor performance (compare for example the green and blue lines for the three q conditions in Figure 8.7).

Table 8.6 BIC value for each performance group

BIC			
	Poor Perf.	**Medium Perf.**	**Good Perf.**
$\boldsymbol{RR_m}$	-0.8968	0.0556	-0.3533
$\boldsymbol{RA}$	-0.7852	-1.1799	-0.8311

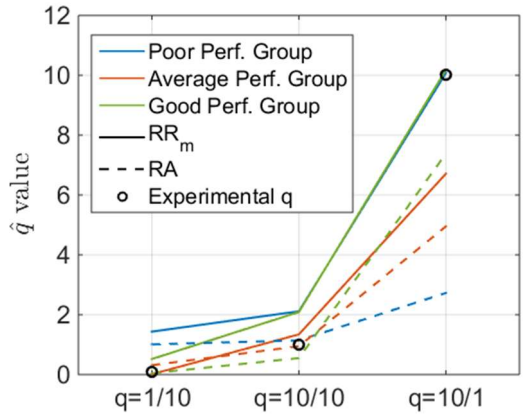


Figure 8.6

Estimated $\hat{q}$ values for each performance group, for the two best fitting decision strategies, RR_m and RA. The figure also shows the experimental q values for comparison (white circles).

8.3.4 Distribution Analysis

The Vincentized distribution of response (Figure 8.8) showed a similar trend found in the previous EXACT paradigm experiment. In particular, the standard deviation appeared to be clearly affected by different λ and different q (for λ: $F_{3,360} = 29.546, p \cong 0$, for q: $F_{2,60} = 17.382, p \cong 0$) . The interaction also resulted to be significant ($F_{6,180} = 6.442, p \cong 0$). The skewness was affected by the $\hat{q}$ conditions ($F_{2,60} = 26.569, p \cong 0$) but not by the λ conditions ($F_{3,90} = 1.883, p = 0.138$). The peculiar shape obtained with some condition (e.g. $q = \frac{10}{10}$ and $\lambda = 0.16$) reflected the difference strategy adopted by different groups. We will compare these distributions with the distributions obtained by the similar Classic RT experiment in the next section. Similar to what observed in previous EXACT experiments, even though most of the distributions presented a slight positive skewness (but negative few cases), this was clearly not as big as the skewness observed in Classic RT task.

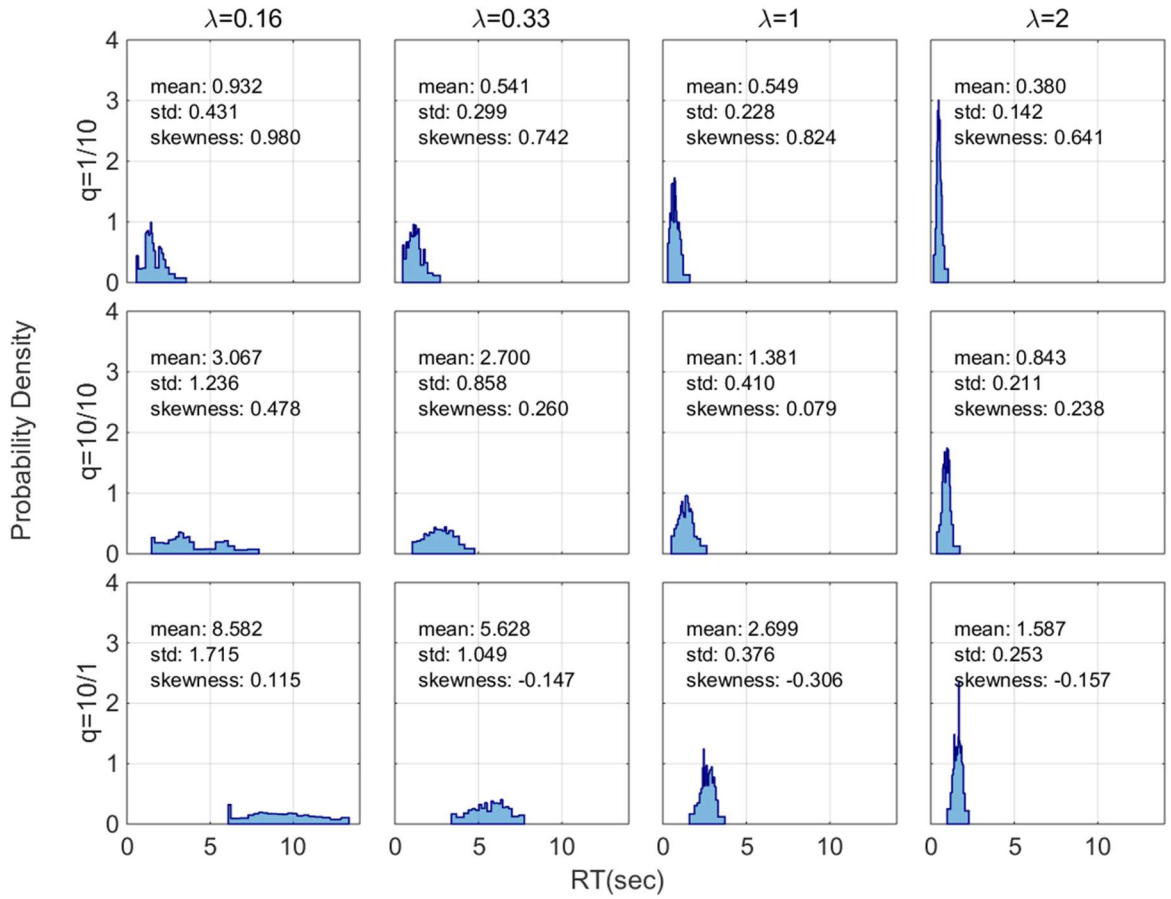


Figure 8.7

Vincentized RT distributions for each condition. As observed before, the standard deviation appears to be consistently affected by the λ value, whereas the skewness is affected only by the q conditions.

Similarly to what found in the previous EXACT experiments, the distribution in the rate domain did not appear to be normal, even after applying the STD-IQR method (Figure 8.9, possible reasons for this are presented in the general conclusion of the previous Chapter, **Section 7.10**). We then investigated the shape of the accuracy distributions, by calculating the aggregated and standardised $ACC(RT)$ for each condition in order to obtain an estimation of the accuracy distribution. Eliminating severely truncated dataset resulted in slightly better distributions, but the difference was not remarkable. Note how from the Kolmogorov-Smirnov test it resulted that 9 out of 12 datasets were not significantly dissimilar from a Normal distribution (at $\alpha = 0.01$). This is consistent to what found in the studies presented in the previous Chapter (see **Section 7.10**).

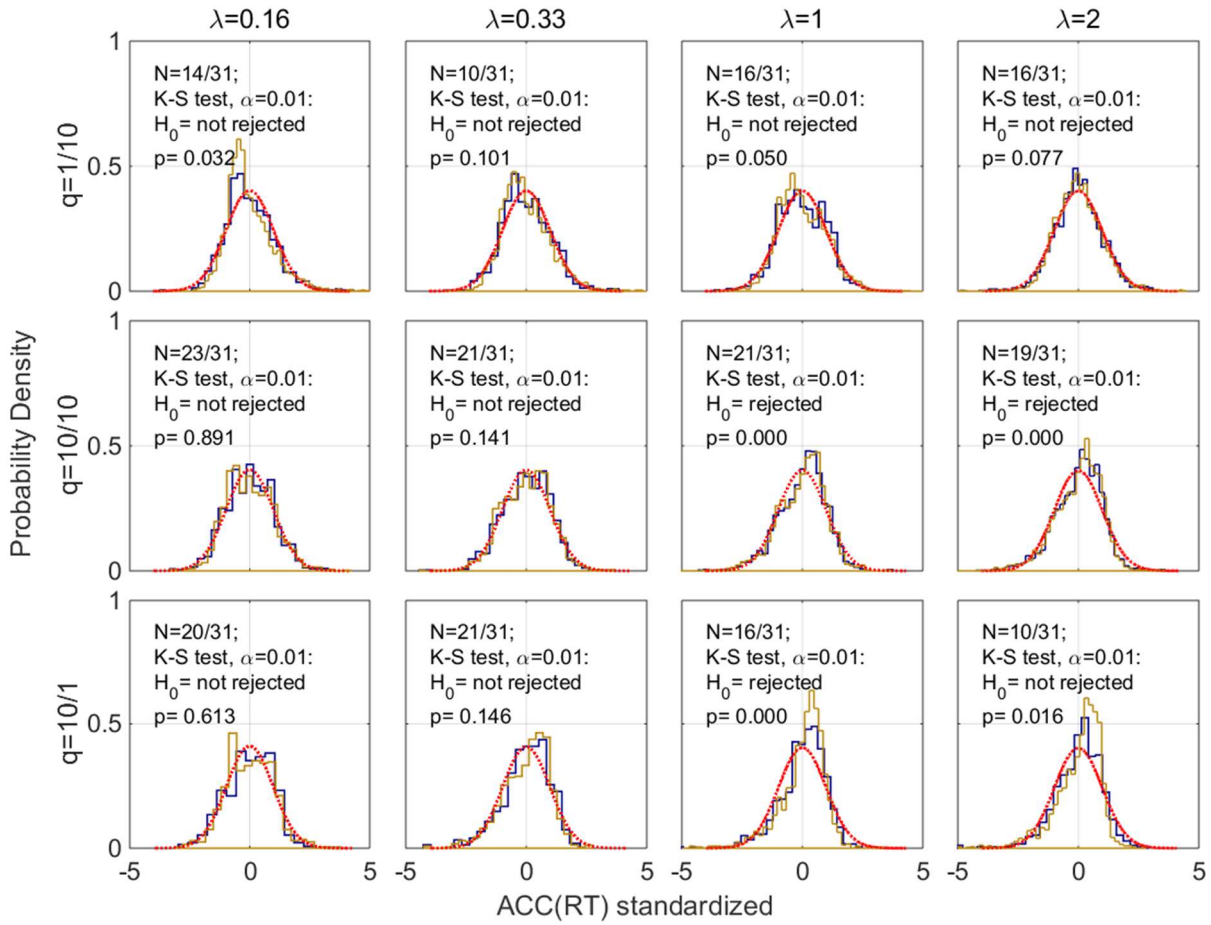


Figure 8.8

Aggregated, standardised accuracy distributions for each condition, with the original dataset (yellow line) and the dataset after excluding severely truncated distributions (blue line). The number of excluding dataset is shown in each panel (N). The Kolmogorov-Smirnov test (on the modified dataset) result is also shown on each panel.

8.4 Discussion

The purpose of this study was to investigate the effect of varying the punishment/reward ratio within the framework of the EXACT Paradigm. First of all, participants were extremely sensitive to change in q, showing that their decision mechanism was actually dependent on the punishment/reward ratio. We also observed that, as with high q the aggregated responses were equivalent to Piéron's Law. Increasing the value of the reward made the responses faster, and the Piéron's Law relationship broke down. The non-monotonic shape observed with α condition (Figure 7.7) was not as clear in this case, at least for the aggregated dataset. However, when dividing across different performance groups, participants in the poor performance group clearly exhibited the non-Piéron's shape that we predicted.

The EXACT Paradigm could be used to examine the $\hat{q}$ variability along a session as done by Balci et al. (2011), but the short length of our condition did not allow us to perform this analysis. Instead, we could only observe that participants in different performance groups used different $\hat{q}$ value, but the analysis was not extremely revealing, as no relevant pattern was found.

An interesting surprise in this experiment was to observe how the RA decision rule was able to account for the data better than RR_m. As we explained, RR_m provided a good fit, but the estimated parameters were very different than the experimental ones, and thus our metric (mWLS) penalized it. Of course, this penalty depends on the different parameters that we used for the kernel function in the mWLS, and also on the weight that we gave to the difference between real and estimated parameters. We interpret this difference with caution, and consider both RR_m and RA as reasonable models in fitting the observations. This was even clearer when we observed the inconsistency across several performance groups, where sometime the RA, and sometime RR_m, appear to provide the preferred fit. To make things more complicated, notice that both RA and RR_m are decision rules depending on the total delay condition which, as we have observed consistently across this work, it does not appear to affect the response. Remember that whereas RR_m is the decision rule that maximizes the reward rate of a condition, RA is a more arbitrary rule used to weight differently speed and accuracy. A general analysis of different decision rules used across EXACT experiments will be provided in the last Chapter (**Section 9.2**).

EXPERIMENT 8.2

In the previous experiment, we investigated the effect of varying the punishment/reward ratio (corresponding to the parameter q for RR_m) across different λ levels, observing that in the condition with higher punishment we clearly obtained a Piéron's-shape curve for mdRT, whereas in the other conditions this was not true. As we discussed in **Section 2.10**, we believe that Piéron's law is not simply a function of the perceptual process used (Stafford & Gurney, 2004) but it is mediated by the decision

rule as well. This is what we have seen within the context of the EXACT paradigm, in which the parameter λ has been interpreted as the stimulus strength. Piéron's law has been observed across myriads of sensory domains and across different experimental designs, so it seems very peculiar that this resides in idiosyncrasy of the used experimental designs.

Even though the EXACT paradigm is a good approximation of a Classic RT experiment, it has some differences (**Section 6.1.6**), and therefore it is important to explore this hypothesis within the context of a more standard Piéron's Law experiment. Furthermore, in this experiment we employed a more classic approach to investigate decision rules used by participants, using quantitative predictions of responses for each decision strategy, based on the DDM (see **Section 2.7**). This approach is limited: avoiding any assumption about the underlying perceptual process was the main point of introducing the EXACT Paradigm. It would be however interesting to compare the results of this study to the results of the previous Chapter, and with the findings from the literature.

Starting from the theoretical paper of optimal decision making (Bogacz et al. ,2006), some authors have tested the predictions generated by different decision rules in the context of DDM. Some of these studies have already been discussed in previous Chapters, for example within the context of changing the delay across trials (D). We are going now to discuss in details all the works with DDM and optimal strategies. The approach in analysing optimal decision consists in comparing the optimal threshold with the empirical one (fitted for each individual participants) and to compare the predicted DT (decision time, function of ACC, or ACC and q, depending on the decision rule) generated by each decision rule. The series of DT for each ACC and q is sometime referred as Optimal Performance Curve (OPC). The optimal threshold a^* is found by using the relationship between DT and ACC (the mathematics is presented in the Method of this Chapter, **Section 8.7**), and requires estimating the parameters of the DDM d (drift rate) and τ (sensory-motor component) and, for some decision rule, it is also function of $D_{tot} = D + \tau$ (D_{tot} is the total non-decision time in a trial, D is the experimental delay in a trial) and $\hat{q}$ (emphasis on accuracy, which needs to be estimated). Once we obtain the optimal threshold a^* we can compare it with the estimated one. We show several OPC in Figure 8.10. Optimal threshold and OPC

are clearly related, as deviation from optimal thresholds corresponds to deviation from the OPC. However, optimizing decision rules generates other qualitative predictions related to accuracy and RT. For example, increasing delay across trial corresponds, for RR_m, RR and RA, to an increase in the threshold, which results in an increase in accuracy and decrease in RT (see **Section 2.8** for an informal discussion, and **Section 8.7** for mathematical details).

Simen et al., (2009) firstly tested the optimal decision boundary predictions given by RR in a task with different reward to stimulus intervals (RSI, first experiment) different bias (second experiment) or different payoff matrix (third experiment). Both experiments consisted of 2AFC with a random-dot-motion task with twelve 4-mins blocks. RT and ACC consistently increased with increasing RSI, but RT was higher than predicted. Higher RSI conditions appeared to affect the threshold value, but this was consistently higher than the optimal threshold. The data from experiment 2 confirmed that participants set their threshold higher than usual, but also obtained qualitative results that were consistent with participants optimizing RR with different bias. Finally, the third experiment explored biased response, and found that participants were able to adjust their behaviour within a single session in order to qualitatively match the prediction of RR. However, in this session as well, participants' estimated threshold were sub-optimally large than predicted ones. Note that this is very different from our manipulation: Simen et al. manipulated the response bias, that is the amount of reward obtained for one or the other correct response, whereas we are focusing on the amount of reward obtained for a correct response and lost for an incorrect one (referring to Table 8.1, they were using a matrix of Type 1, whereas we used a matrix of Type 2).

Bogacz et al., (2010) analysed three different decision rules: RR_m, RA and RR, by using two experimental designs: random dot motion task and asterisk counting. For both experiments, the RSI was varied across sessions. The main finding was that the threshold consistently increased with increasing RSI (resulting in an increase in accuracy and RT). However, it was found that participants on average set their threshold much higher than optimum predicted by RR and, similarly, their DT was higher than the one predicted by OPC. Using RA and RR_m decision strategies resulted in more precise

better predictions in terms of DT. This was especially interesting considering that, in the task, incorrect responses were not explicitly penalized. Both RA and RR_m provided a better fit than RR for the three performance groups in which the sample was divided, but RA provided a much better results the RR_m for the bottom 10% participants.

These results, showing how participants set their threshold sub-optimally large, were partly backed up by Balci et al. (2011) which, in a random-dot motion experiment with different coherence levels, compared RR_m and RR. It was found that, while participants appeared to favour accuracy at the beginning of the task (with RR_m providing the best fit), they decreased or eliminated this emphasis with training, therefore favouring RR after many sessions (recall that RR corresponds to RR_m with $q = 0$). This effect interacted with signal quality.

So, whereas the results from Balci et al. suggests that RR_m (with a decreasing and vanishing q) may fit better, Bogacz et al. (2010) found that RA provided the best fit and Simen et al. found a good fit between prediction and observation by only using RR. These mixed results suggest that the decision strategy may interact in complicated way with experimental setups and designs.

Summarizing, the optimization approach within the DDM framework has been investigated by using several approaches: changing the alternative bias (Simen et al. 2009), the RSI across trial (Bogacz et al., 2010), the stimulus strength (by changing the motor coherence in a random-dot-motion task, Bogacz et al., 2010; Balci et al., 2011; Simen et al., 2009) and the reward bias (Simen et al., 2009). However, investigating the outcome for different punishment/reward ratio has yet to be done. The main reason is that the other manipulation clearly corresponded to manipulating a parameter within the theoretical model of DDM and decision strategies, where, as discussed in **Section 8.1**, the parameter q is not clearly mapped to the punishment/reward ratio for all decision strategies, but it corresponds to a general "emphasis on accuracy".

8.5 Methods

Participants. 20 participants (10 males, 10 females) took part in the experiment. All participants had normal or corrected-to-normal vision and no known neurological condition. The study received ethical approval from the local ethics committee.

Material. The experiment was a 2AFC task in which participants had to identify in which side of the monitor a Gabor Patch appeared. The study was conducted under constant levels of illumination. The stimuli were presented binocularly on a computer monitor (Sony GDM-F520) using the software E-Prime (version 2.0.10). Participants were positioned at 63 cm from the monitor. The stimuli were Gabor patches presented on the left or on the right of the centre screen. Each patch was a vertical sine-wave grating multiplied by a Gaussian windows function ($\sigma = 0.06°$), trimmed for values lower than 0.005. The spatial frequency was 4 c/deg. The stimuli were presented within a grey square window (10 cd/m^2) of 5 degrees/cm. The rest of the monitor screen was kept dark (~0 cd/m^2). We used the 4 different contrast levels: 0.035, 0.050 0.1 0.2. Participants obtained X points for a correct response (reward) and lost points Y for an incorrect response (punishment). The amount of points lost or earned depended on the q condition, which specified the punishment/reward ratio. We used 4 q conditions: $\frac{1}{30}, \frac{1}{10}, \frac{10}{10}, \frac{10}{1}$ (from the more rewarding, to the most punishing).

Procedure. The experiment was presented as a point game, and participants were asked to earn as many points as possible. After instruction and a short training session (30 seconds for each contrast level, with $q = 10/10$) the main session started. A black dot presented in the middle of the grey square window was used as a warning signal, and was presented for 0.75 second (foreperiod, FP), after which a vertical line appeared on the screen and, at the same time, a Gabor Patch appeared on one of the side of the line, within the grey square. Participants were instructed to press the key “q” or “o” depending on the side of appearance of the stimulus. After responding, the patch and the vertical line disappeared and the screen showed the score for the trial (eg. “+10”, “-30”) and the total score for that session for 0.75 seconds (reward-to-warning signal interval, RWI). There was no audio feedback. Each trial lasted 2 minutes regardless of participant’s performance. For each q condition, we ran all the contrast levels

(blocked and in randomized order) before proceeding to the next q condition. We always ran $q = \frac{10}{10}$ as the first q condition. At the end of the experiment, the points were exchanged for money, with the exchange rate of 60 points=1 penny. This unfavourable exchange rate is balanced by the fact that, across the whole experiment, participants obtained thousands of points. See Figure 8.10 for an illustration of the flow of the task.

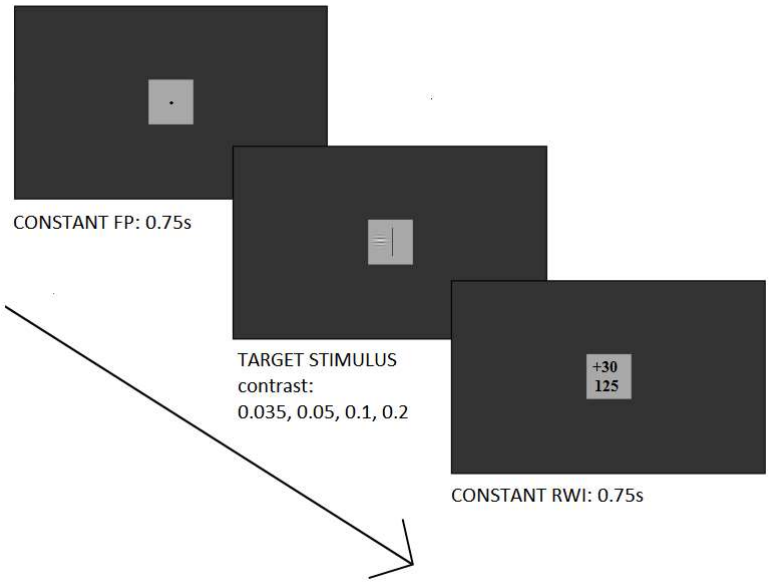


Figure 8.9

The flow of time for the task.

Analysis. We fitted the distribution uniquely to the Pure DDM. We have seen in the previous Classic RT Experiment (Chapter 4 and 5) that the Pure DDM provided a reasonably fit by using also fewer parameters, which results in a lower AIC. Although the Extended DDM may provide a better fit in some situations, the analysis in terms of optimal decision has not been found to be amenable to formal analysis (Bogacz et al., 2006), and numerical simulations are sometimes used. Noisy estimation of parameters and parameter inflation makes this approach difficult (Simen et al., 2009). There is also another reason why we do not apply the Extended DDM to quantitative optimal performance analysis: it has been shown (Drugowitsch et al., 2012) that the optimum boundary for a DDM with noisy drift rate (which is used by the Extended DDM) is not a constant one, as usually assumed, but a non-linear

decreasing one (see also Moran, 2014). It is, therefore, misleading to conduct optimality analysis with the Extended DDM with constant boundaries, and a moving boundary model should be employed, which is however beyond the scope of the present work.

8.5.1 Optimal Threshold and Optimal Performance Curve

The relationship between decision strategies and DDM parameter is presented in Chapter 2, **Section 2. 8**. By assuming the Pure DDM, an optimum threshold and the optimum DT can be found for each decision rule. Note that in Bogacz et al. (2006), where this relationship was presented for the first time, a was equal to the distance between the starting point and the threshold, whereas in our work a is the distance between the two boundaries. To switch from one definition to another we used $a_B = a(1-b)$, where a_B is the threshold distance to starting point (as defined in Bogacz et al., 2006), a is the distance between the two thresholds, b is the bias (note that in the present work, the bias parameter is always set at 0.5, meaning that no bias is assumed between the two alternatives). This allow us to define the accuracy function as

$$ACC = 1 - \left(\frac{1}{1+\exp\left(\frac{2da(1-b)}{s^2}\right)}\right)$$

and

$$DT = \frac{a(1-b)}{d}\tanh\left(\frac{da(1-b)}{s^2}\right)$$

The optimum threshold for the decisions strategies can be easily found by substituting $ACC(d,a)$ and DT into the decision strategy equation and calculate the derivative with respect to a (Bogacz et al., 2006; Bogacz et al., 2010). This usually results in transcendental equation which must be estimated numerically.

Reward Rate. For RR, the optimum normalized threshold a^* must satisfy

$$\exp\left(\frac{2da(1-b)}{s^2}\right) - 1 = \frac{2d^2}{s^2}\left(D + t_0 - \frac{a(1-b)}{d}\right)$$

Which solution can be computed numerically. The family of optimum thresholds for different d reaches its maximum at the point where $a(1-b)d = s^2$ (s is set to 1 in this work).

By using the relationship between ACC, d, a and b above we can compute the Optimum Performance Curve which describes the relationship between ACC and RT:

$$RT^*_{RR} = (D+\tau)\left(\frac{1}{(1-ACC)\log\left(\frac{ACC}{1-ACC}\right)} + \left(\frac{1}{2ACC-1}\right)\right)^{-1} + \tau$$

(slightly rearranged from Bogacz et al., 2006). Recall that DT can be easily found: $DT = RT - \tau$. This shape is presented in Figure 8.11. Note that for computing this curve we only need to estimate the τ parameter from the DDM, whereas ACC is behaviourally observed and D is an experimental parameter. This is the decision rule we were less focused on, as it does not take into account the punishment/reward ratio and it seemed unlikely to account for tasks with different payoff matrices. Note that we can rearrange the terms so that $\frac{DT}{D_{tot}} = f(ACC)$, where $DT = (RT - \tau)$. This implies that, having computed DT for each individual participant, it is possible to pool data across participants and task conditions, irrespectively of task difficulty, timing, or individual differences.

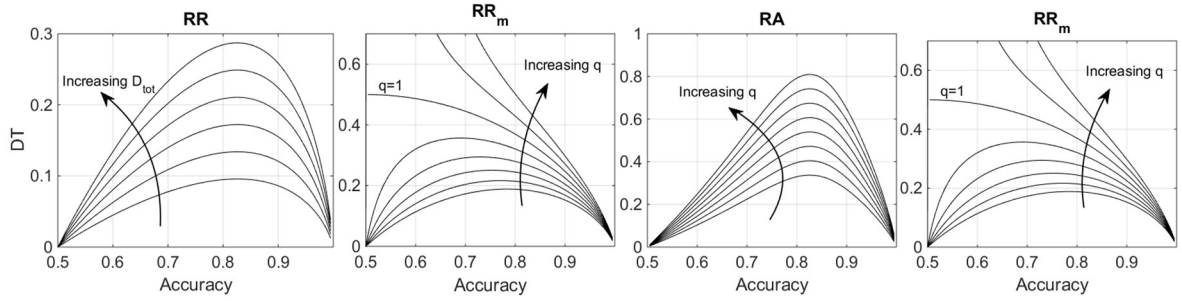


Figure 8.10

Optimum decision time ($DT*$) curves given a certain value of accuracy, for each decision strategy.

Bayes Risk. The optimum threshold for the BR decision rule has been found by Edwards (1954) to be the value that satisfies the expression:

$$\frac{q2d^2}{s^2} - \frac{4da(1-b)}{s^2} + \exp\left(-\frac{2da(1-b)}{s^2}\right) - \exp\left(\frac{2da(1-b)}{s^2}\right) = 0$$

Which has a unique solution (Bogacz et al., 2006). With $q = 0$ (no emphasis on accuracy, only on DT) the optimum threshold is zero. The optimum response time is found to be

$$RT^*_{BR} = q\left(\frac{(2ACC-1)\log\left(\frac{ACC}{1-ACC}\right)}{2\log\left(\frac{ACC}{1-ACC}\right) - \frac{1}{ACC} + \frac{1}{1-ACC}}\right) + \tau$$

(slightly rearranged from Bogacz et al., 2006). Note that the optimum response is not function of the total trial time. Is it, however, function of the parameter q, which establish the emphasis on accuracy. The curve for DT^*_{BR} (equal to $RT^*_{BR} - \tau$, as presented in Bogacz et al., 2006) for several parameter sets are shown in Figure 8.10. For these and the following functions, note that we can rearrange the terms so that $\frac{DT}{D_{tot}} = f(ACC, q)$ (as in Bogacz et al., 2006).

Reward Rate Modified and Reward\Accuracy. Similarly, the optimum threshold for RR_m has to satisfy the following equality:

$$\frac{\frac{2dE}{s^2(E+1)^2} + \frac{2dqE}{s^2(E+1)^2}}{(D+\tau) + \frac{aT}{d}} + \frac{\left(\frac{T}{d} - \frac{a(1-b)(T^2-1)}{s^2}\right)\left(\frac{1}{E+1} + \frac{q}{E+1} - 1\right)}{\left((D+t_0) + \frac{aT(1-b)}{d}\right)^2} = 0$$

Where $E = exp\left(\frac{2ad(1-b)}{s^2}\right)$ and $T = tanh\left(\frac{ad(1-b)}{s^2}\right)$, which results on an optimum response equal to:

$$RT^*_{RR_m} = (D+\tau)(1+q)\left(\frac{\frac{1}{1-ACC} - \frac{q}{ACC}}{log\left(\frac{AC}{1-ACC}\right)} + \frac{1-q}{2ACC-1}\right)^{-1} + \tau$$

For Reward\Accuracy, the optimum threshold has to satisfy

$$\frac{\left(\frac{1}{E+1}-1\right)\left(\frac{T}{d}-\frac{\left(a(1-b)(T^2-1)\right)}{s^2}\right)}{F^2}+\frac{2dE}{s^2(E+1)^2F}+\frac{2dqE}{(D+t_0)s^2(E+1)^2}=0$$

With E and T defined as above, $F=(D+\tau)+\frac{a(1-b)T}{d}$, and the optimum response is equal to

$$RT^*_{RA}=(D+t_0)\frac{G-2q\sqrt{G^2-4q(G+1)}}{2q}+\tau$$

With $G=\frac{1}{(1-ACC)\log\left(\frac{ACC}{1-ACC}\right)}+\frac{1}{2ACC-1}$. Note that both these decision rules are function of the delay across trial D and the emphasis on accuracy parameter q, the first being and experimental parameter, whereas the latter must be fitted to the observations. Note that the interpretation of these parameters is the same as the one illustrated in Chapter 6: q is equivalent to the punishment/reward ratio for RR_m and to the weight of accuracy (scaled for the total delay $\tau+D$) against the weight for RR for the RA decision rule. It is very common the literature to consider only values of q lower than 1, as for values higher than 1 and for the decision strategy RR_m , not all values of ACC are well defined (Zackenhouse et al., 2009; Bogacz et al., 2010). However, we have observed as participants are generally loss aversive, and thus weight accuracy much more than speed. Plus, the estimated values of q in the EXACT experiment were almost uniquely higher than 1. Thus it seems unreasonable to consider only values lower than 1. By considering any $q>1$ we are saying that, sometime, participants will never reach some low level of accuracy (because, for example, the penalty is very high) when they respond optimally (see Figure 8.10). In fitting, we have the problem of deciding how to fit the observed DT for ACC that do not have any associate DT. We decide to give the high penalty by pretending that those responses were 10 seconds afar from optimal response.

8.6 Results

8.6.1 Signal Analysis and Response to Feedback

We firstly analysed the result in terms of signal of response, by using the STARS method in the same way as done with the EXACT Paradigm. We found that the average number of shifts in the signal was lower than 1, and the magnitude of the shifting was also, on average, lower than 1. The response signals in this experiment were slightly more stable than the one in the EXACT Paradigm, where there were on average 2 shifts per signal, with approximately the same magnitude. This increased stability across condition is confirmed by the analysis of aggregated patterns of response for each q condition. In the EXACT experiment we found several trends, depending on the q condition, whereas here we found no significant trend in any condition. Overall, this confirmed what we expected from previous results: participants in Classic RT task do not appear to adjust significantly their response across condition. Is it however possible that participant did not have enough time to adjust their response, as session length was kept short (2 minutes). In Balci et al., (2011), where participants performed 13 5-minutes session of dot-motion discrimination task, RT decreased across session, which was interpreted as a steady decrease in the emphasis on accuracy.

The results in terms of median of median RT (mdRT) are shown in Figure 8.12. We ran a 2-way repeated measure ANOVA and found that manipulating contrast significantly affected median rate ($F_{3,57} = 41.160, p \cong 0$), but manipulating the punishment/reward ratio q did not reach significance ($F_{3,57} = 3.475, p = 0.022$). The interaction between the two factors was not significant ($F_{9,171} = 1.30, p = 0.240$). Post-hoc pairwise comparisons with Bonferroni correction revealed that each contrast condition was significantly different from each other ($p < 0.005$), a part from the lowest with the second to lowest contrast. A similar pattern was found for the effect on accuracy (contrast: $F_{3,57} = 66.677, p < 0.0001$, q: $F_{3,57} = 1.562, p = 0.209$).

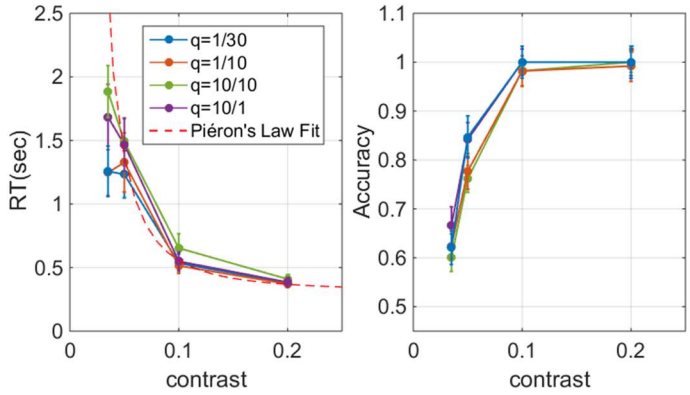


Figure 8.11

(**Left panel**) Median of median of responses (circles connected with a line). One Piéron's function is fitted to all q conditions (red dashed line). The Piéron's function is fitted only from the second contrast condition onwards. (**Right panel**) Median of estimated accuracy for each participant.

Even though the results are not significant, it appears from the plot that for some conditions the mdRT followed a Piéron's shape, and for the rewarding condition it followed a non-Piéron's shape (see the non-monotonic shape obtained for the low contrast levels). Comparing this results with the results from the previous study (Figure 8.4), it appeared that here participants were not really sensitive to changing in the payoff matrix.

We fitted Piéron's Law to the data. However, since some of the q conditions appeared to have an inflection point with the lowers contrast condition, we fitted Piéron's Law only on the results with the contrast equal to 0.035 0.05, and 0.1. We allowed the parameter α, β or γ or any combination of these parameters to be free across q conditions, and we fitted the function by minimizing the weighted least square. The preferred model was calculated by applying a BIC criterion for least square. The preferred model resulted to be the one with no free parameters across q conditions (BIC=-4.0328, $df = 3$), confirming the lack of effect of payoff conditions. The one-line fit is shown in Figure 8.12 (red dashed line).

By visual inspection of median RT for individual participants, we noticed that several participants median RT had a non-Piéron's shape, with the inflexion point at the second contrast level. For most of them the difference across q condition did not seem remarkable, so that when the peculiar non-Piéron's

shape was present, it was present for all q conditions. We can better understand the result by grouping participant according to their performance. We can see a clear difference between mdRT across difference performance group in Figure 8.13.

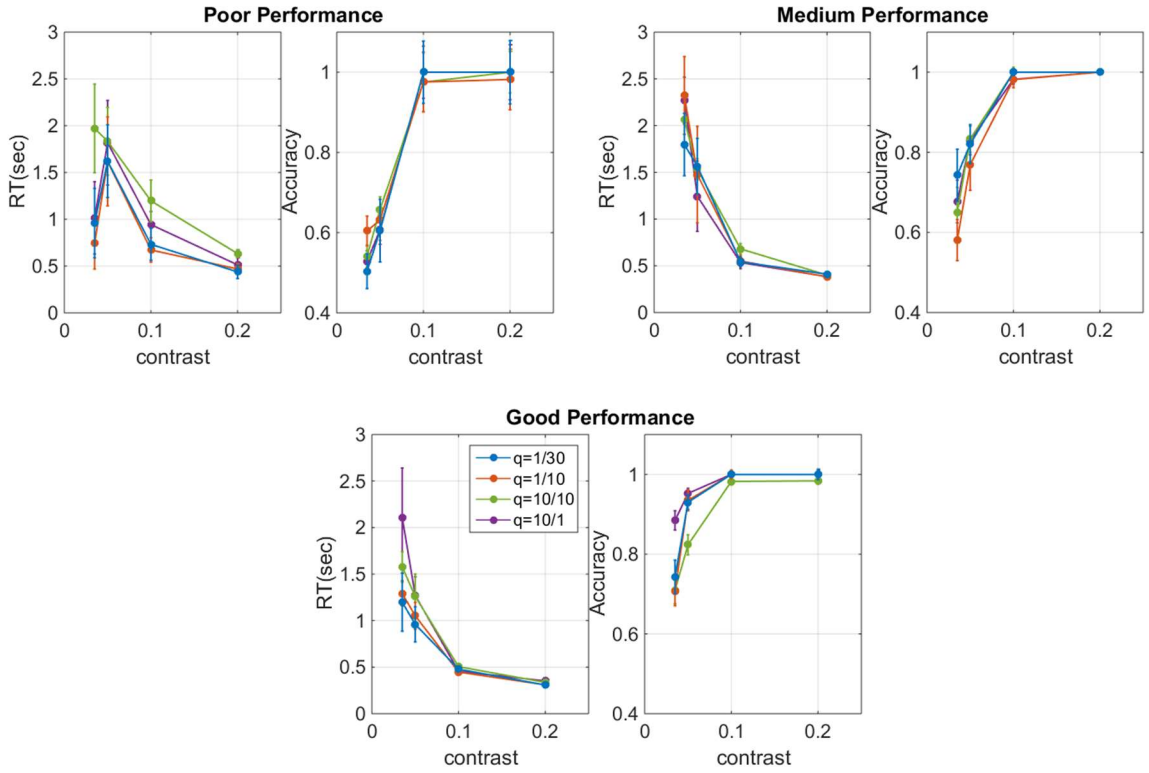


Figure 8.12

Responses and accuracy for each performance group. Each bar indicates one standard deviation.

Responses for the participants in the poor performance group clearly followed the non-Piéron's shape, whereas participants in the other two groups followed the more classical Piéron's shape. The effect of q appeared to be small and was non-significant for any of the performance group. Moreover, participants in the good performance group were faster than participant in the medium performance group. Why only participants in the poor performance group generated the non-Piéron's shape, regardless of the contrast level? In terms of decision rules, participants in the poor performance group may be wrongly estimating $q, d,$ or may have a slowly-increasing speed-accuracy function than the participants in the other two groups because, for example, they may have poorer vision. As we explored in the previous Chapters, changing q, d or α can effectively change the shape of the RT^*. However,

when the q and d is increased (or α is decreased) the result is to switch from a non-Piéron's shape to a Piéron's shape, but then keep increasing q or d (or keep decreasing α) would result in a further increase in RT, not in a decrease like we observe between the good and the average performance group. Therefore, a change in $q, d,$ or α, cannot explain the difference between the performance groups. A change in λ, however, can: assuming that participants in the poor performance group had poorer vision is the only way to obtain the pattern within our framework. The exact relationship can be obtained only by knowing the real speed-accuracy function. However, an intuitive explanation, simulated by using the exponential $ACC(RT)$ and for 4 q values (0.8, 0.85, 0.9, 0.95) and $D_{tot} = 1.5s$ is provided in Figure 8.14. We assumed that participants can be divided in 3 groups based on their visual acuity. Participants with better vision can extract more information from the stimulus, and thus may have a subjectively higher stimulus strength than the other groups: this is associated with lower RT and a Piéron's shape RT curve. On the other hand, participants with poor vision will experience a weaker stimulus strength, and thus they will be on the left side of the decision curve, within the inflection point, generating a non-Piéron's shape.

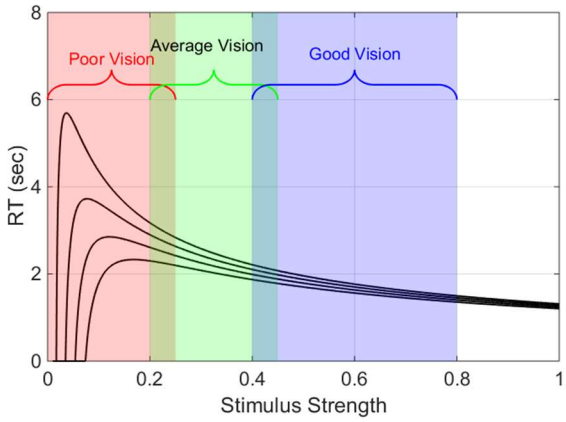


Figure 8.13

Illustration of a model explaining the different RT curves for difference performance groups. It shows a hypothetical optimum response (RT^*) curve and three possible ranges of vision. Participants with a poor vision (red area) will perceive a weaker stimulus strength than participants with a good vision (blue area).

8.6.2 Distribution and Model Fitting

The Vincentized distributions are shown in Figure 8.14 with the respective mean, standard deviation and skewness averaged across participants for each condition. As expected, standard deviation tended to strongly decrease with increasing in contrast (this appeared to be significant, $F_{3,57} = 55.519, p \cong 0$), whereas skeweness appeared to increase (this was also significant, $F_{3,57} = 7.610, p \cong 0$). The effect of q on standard deviation and skewness was, however, not significant. ($F_{3,57} = 3.19, p = 0.03$; $F_{3,57} = 0.319, p = 0.811$). The effect of contrast on the standard deviation was consistent with the effect of λ found in the EXACT paradigm. As for the skewness, in the previous works λ was not found to be significantly connected with the skewness distribution, whereas here the contrast appeared to have an increasing effect on it. One reason could relate to the fact that λ was fixed across trials, whereas in a Classic RT the noisy stimulus strength will produced more skewed distributions.

The analyses of optimality performance with aggregated data has been performed and discussed by Simen et al. (2009). We fitted the aggregated (with the Vincentizing method) data with 10 quantiles. We used aggregated participants (regardless of the advantage given by analysing individual results) because for some participants the responses for an individual session were as low as 25, and the average response number was 50, too low for obtaining a reliable estimate of parameters. Note that, as in the previous Chapter, we used the Quantile Maximum Estimate Method, which scales automatically for the error proportions (see **Section 2.3**). We allowed the parameter d to vary across contrast conditions. We then compared four different models: with a or τ allowed to change across individual conditions, with both a and τ allowed to change across q conditions, and with no parameter (but d) allowed to change across conditions (see Table 8.6).

Table 8.7 Model variants used for the DDM

Variant 1	Free τ for q conditions
Variant 2	**Free a for q conditions**
Variant 3	Free a and τ for q conditions
Variant 4	Nothing free

In this way, we can estimate whether there was an effect on the distribution of the q conditions, and what parameter did it affect. Interestingly, in spite of the lack of significance between q conditions and mdRT, the best model was given by the one with free a across different q conditions, with slightly increasing threshold values with increasing punishment (increasing q), as shown in Table 8.7. This implies two things: 1) changing q condition may not have a big effect on mdRT, but it affected distributions (and proportion of errors) enough to be significant when fitting the pure DDM; 2) changing the q condition had an effect on the speed accuracy trade-off, as the effect on threshold suggested. The effect on d was, as expected, increasing with increasing contrast. However, the estimated parameters also revealed that the changing in q was not particularly remarkable (see Table 8.8), and not always increasing with increasing q. The resulting fitting is shown on the distribution in Figure 8.15.

Table 8.8 Model fitting results

Pure DDM		
	AIC	**df**
Variant 1	47160.11	9
Variant 2	**43192.81**	**9**
Variant 3	49188.8	12
Variant 4	44488.13	6

Table 8.9 Best Model's parameters

Variant 2 (Pure DDM)	
AIC	43192.81
$\boldsymbol{\Sigma\,\mathcal{L}}$	43174
$\boldsymbol{a_{q=\frac{1}{30}}}$	1.932
$\boldsymbol{a_{q=\frac{1}{10}}}$	2.241
$\boldsymbol{a_{q=\frac{10}{10}}}$	2.236
$\boldsymbol{a_{q=\frac{10}{1}}}$	2.453
$\boldsymbol{\tau}$	0.196
$\boldsymbol{d_{c=0.035}}$	0.221
$\boldsymbol{d_{c=0.05}}$	0.668
$\boldsymbol{d_{c=0.1}}$	2.658
$\boldsymbol{d_{c=0.2}}$	4.769

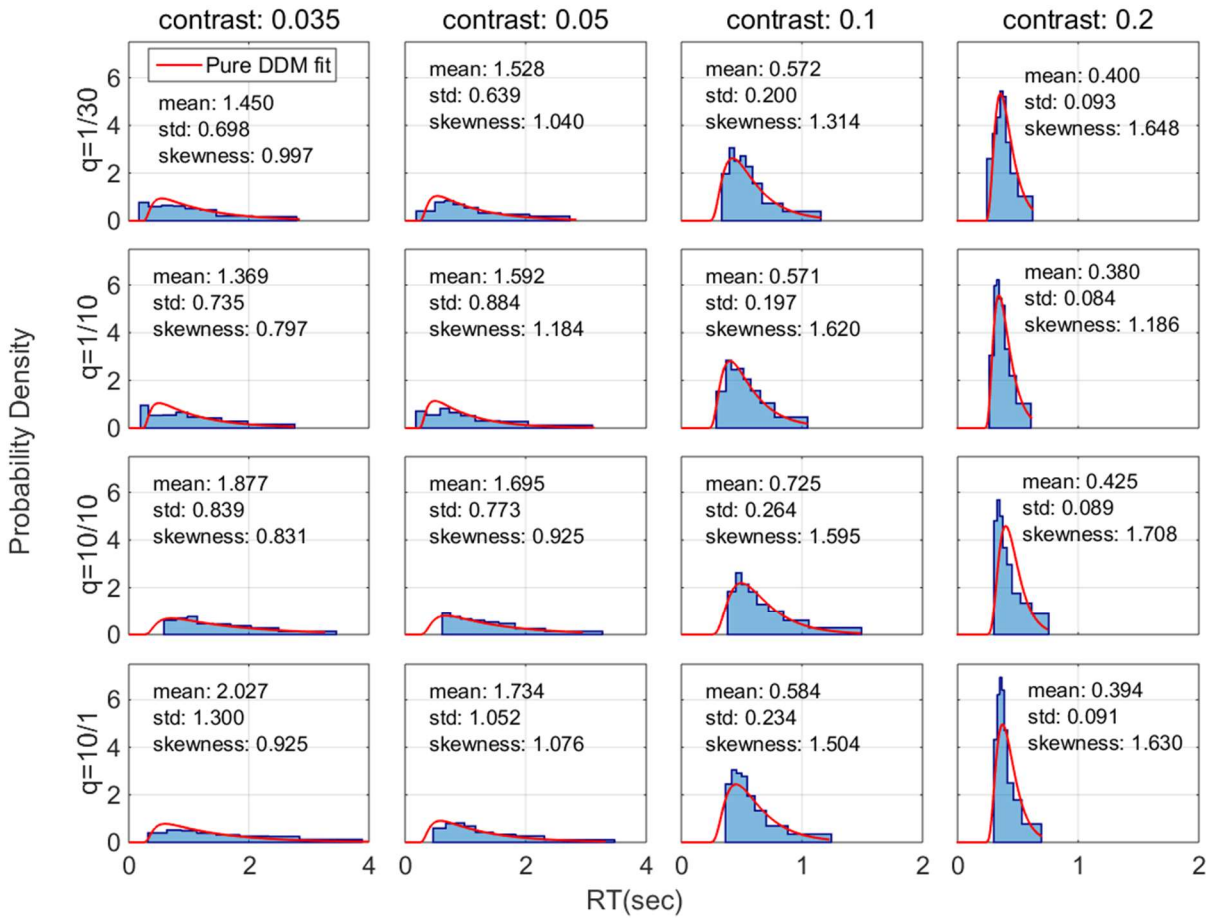


Figure 8.14

Vincentized RT distributions for each condition, with the obtained Pure DDM fitting shown on each panel.

8.6.3 Comparison with Optimum Response

We firstly calculated DT in two different ways. Firstly, we assumed that there was no relevant difference across q conditions. We pooled all the participants' mean response together, regardless of q or contrast condition, and divided the possible range of ACC (0.5 to 1) in 10 equal intervals. For each interval, we calculated the mean normalized DT ($DT = RT - \tau$, where τ is estimated to be 0.196s). For random chance, participants sometime obtained an Accuracy that was lower than 0.5 in some conditions. We excluded these values from the analysis (7% of total data). The result is shown in Figure 8.16 (the bar corresponds to one standard deviation).

To fit the data we used the same procedure used in the previous Chapter to fit mdRT to optimal response curve for each decision strategy. As before, we assumed that participants' estimation of the delay across trial $D_{tot} = D + \tau$ (with D equal to 1.5, that is the constant RWI equal to 0.75 plus the constant FP equal to 0.75) may have slightly deviatet from the real estimation, and the weighted least square measure was modified accordingly to take this into account. We used the same modified WLS employed in the previous Chapter. We found that the weight parameter ω of 1 was too permissive, allowing deviation from the real D parameter as high as 10 seconds, and thus we increased it to 1.5, which resulted in deviation in the range of 1 second. Note that we will refer to the estimated D_{tot} value as $\widehat{D}_{tot}$. For RR, this was the only parameter to fit. For RR_m, RA and BR we also fitted the q parameter. For this initial set of analyses the results are shown in Figure 8.16 (red lines). Resulting BIC values for each decision strategy is shown on Table 8.9, whereas Table 8.10 shows the estimated parameters for each decision strategy. The results were, generally, unsatisfying, as none of the decision rules' predictions resembled the observed shape.

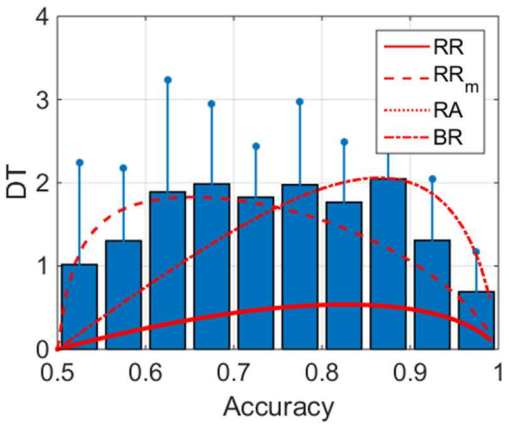


Figure 8.15

Mean DT across participants binned for different levels of accuracy. Bars correspond to one standard deviation. The red lines correspond to the resulting fitting for the four decision strategies.

Table 8.10 Results for the decision rules fitting

	BIC	df
$\boldsymbol{RR}$	1.41078	1
$\boldsymbol{RR_m}$	-0.66781	**2**
$\boldsymbol{RA}$	0.06340	2
$\boldsymbol{BR}$	-0.36914	2

Table 8.11 Estimated parameters

	$\boldsymbol{RR}$	$\boldsymbol{RR_m}$	$\boldsymbol{RA}$	$\boldsymbol{BR}$
$\widehat{\boldsymbol{q}}$	-	0.954	1.056	15.135
$\widehat{\boldsymbol{D}}_{\boldsymbol{tot}}$	2.806	2.243	2.199	-

We supposed that this may have been due to our method of collapsing participants that used different strategies and thus tried other grouping methods. Firstly, we repeated the same analysis by separating data according to the q conditions, but there was not much difference between DT for different q condition (as observed before).

Similar to what done in the previous Chapter and in Bogacz et al. (2010), we separated the participants according to different performance. We fitted the DDM to each performance group (we excluded the errors in these cases, as their number was too small). The results were difficult to interpret: for the poor performance group, the best DDM variant was the one with varying threshold separation a across different q conditions, which meant that these participants were actually adjusting their response according to how much reward and punishment they were expecting to get. Surprisingly, for the other two performance groups, the three variants gave roughly the same AIC results, meaning that there was not much difference in the fitting power of the three models, and participants may have slightly adjusted τ, a, or none of them (the model with both τ and a free to move across q condition resulted in a very high AIC value due to the higher number of parameters). The estimated parameters for these models were also roughly the same.

The interpretation of these different τ values was most problematic. For the poor performance group, τ (for the model with free a across q conditions) was equal to 0.05s. This must not be particularly surprising, especially considering Figure 8.13, which hints to the fact that participants in the poor performance group were using the constant trial delay ($FP + RWI$) to estimate the onset of the stimulus, and responded in a non-integrative way (that is, without accumulating any information). For the other two performance groups, τ slightly varied across different model variants, and even with free τ across q conditions the values oscillated within a range of only 40 ms. We took the average of these measurements resulting on an estimated τ of 0.16 seconds for both medium and good performance groups. The resulting DT for each performance group is shown in Figure 8.17, and the AIC fitting in Table 8.8, showing RR_m is the preferred model for medium and good performance, whereas RA is the preferred model for the poor performance group. Some cautions must be used in interpreting these results: the DT did not resemble the inverted U-curve shape predicted by all decision rules. For the poor and medium performance groups we observed a decrease on the extreme values of accuracy, but overall the responses were very noisy and, as a result, the fitting of the decision strategy was not really

revealing. The reason for this generally unsatisfying fitting will be investigated in the Discussion Section.

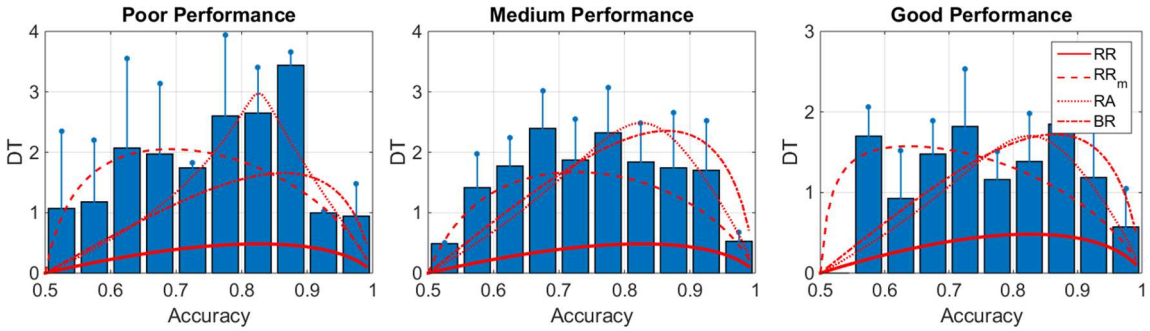


Figure 8.16

Mean DT curves for each performance group, with the corresponding best DT curve fitting for each decision strategies. None of the observed results appears to be close to the expected optimum DT.

Table 8.12 Fitting results for each performance group

	BIC values		
	Poor P.	**Medium P.**	**Good P.**
$\boldsymbol{RR}$	5.04881	2.510413	2.355007
$\boldsymbol{RR_m}$	1.566364	0.089207	0.7473
$\boldsymbol{RA}$	1.150535	1.406435	0.824595
$\boldsymbol{BR}$	1.822022	1.172989	0.183615

8.6.4 Rate Distribution Analysis

The rate domain distributions are shown in Figure 8.18. We already analysed the Rate Domain Distributions for Classic RT experiments in Chapter 4 and 5 and found that in most cases it followed a Normal Distribution. We applied the same analysis to this dataset. The blue lines represent original distributions, whereas the yellow lines represent distributions after elimination responses that did not belong to the original distribution (we used the STD-IQR method, as explained in **Section 3.3**). However, the number of cut data points did not appear to produce a relevant difference for most distributions. Compare, for example, this plot with that in Chapter 4 (Figure 4.7), where the distributions were remarkably "peaked" due to anticipatory responses produced for the Simple RT task. The Kolmogorov-Smirnov test ($\alpha = 0.01$) applied to the modified dataset, revealed that for 11 distributions

out of 15 the normality hypothesis could not be rejected. Note that the distributions that most deviates from Normality were those with very low contrast, regardless of the q condition. The contrast level used here is also much lower than the one used in previous Chapters, and the distributions with the lowest contrast levels are also those with the most degenerated shape. As shown above (Figure 8.13), several participants had very fast responses with low stimulus contrast, which may hint to the fact that they were basing their response mostly to the FP length (so to respond soon after the stimuli appeared), and thus generating distribution that are a mixture of two different processes, similar to what we observed in Chapter 4. However, this is only a speculative hypothesis, and it will be discussed in the General Discussion.

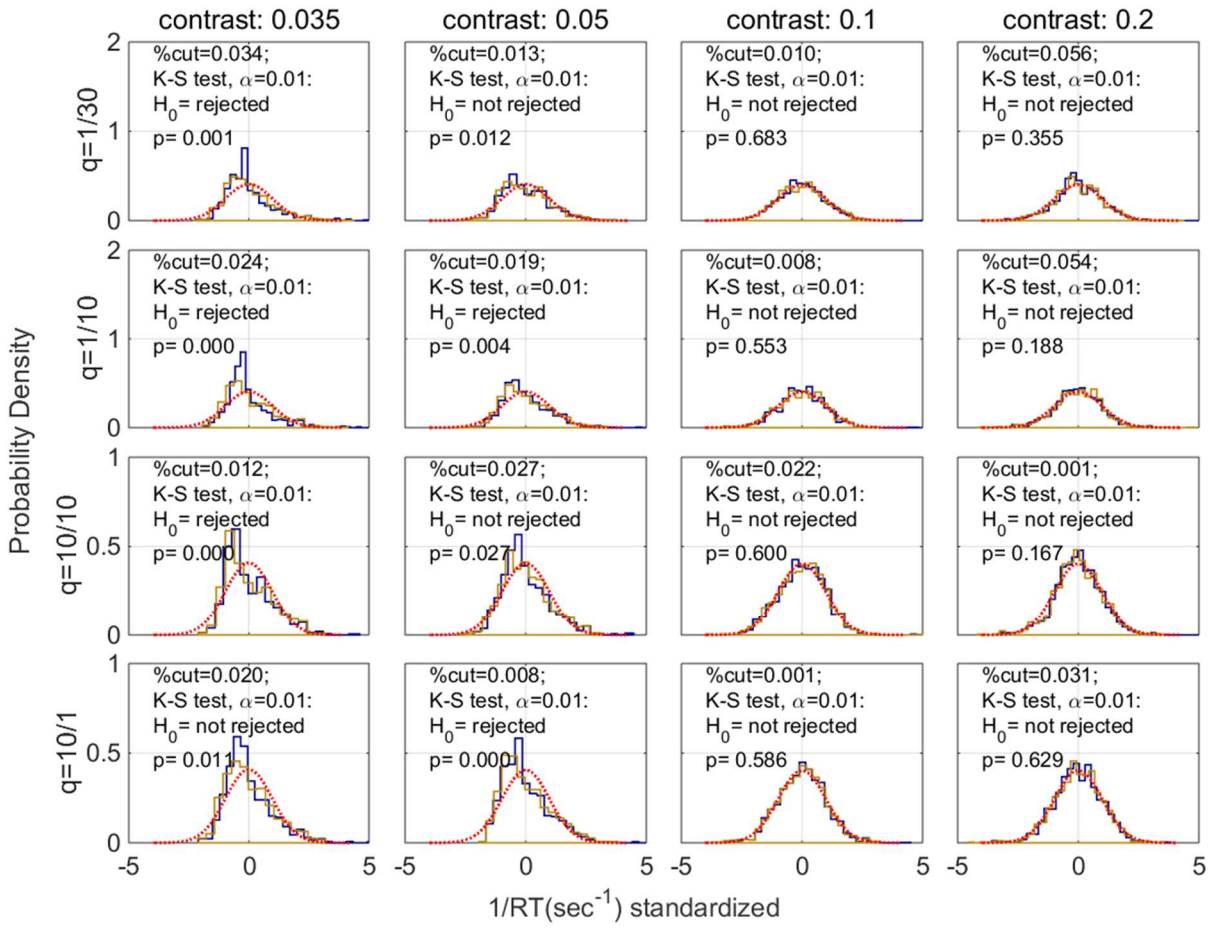


Figure 8.17

Aggregated and standardised accuracy distribution, before (blue) and after (yellow) applying the STD-IQR method. The red-dashed lines correspond to a Standard Normal Distribution.

8.7 Discussion

The double aim of this work was to verify that non-Piéron's shape was actually possible within the context of a Classic RT experiment; to investigate the effect of different punishment/reward ratio; to analyse the response in terms of optimality by assuming the Pure DDM model. The non- Piéron's shape was effectively observed, but in a different context than predicted. We expected to observe Piéron's shape for three of the four q conditions (with balanced, high, and very high punishment), and a non-Piéron's shape for the condition with higher reward, whereas we observed it for the poor performance group, across all q conditions. Some participants were effectively using a non-integrative strategy (that is, responding randomly as stimulus onset, without accumulating any information), apparently in an unrelated way with the q condition, but likely due to their poor vision (that is, for them the lowest condition was difficult enough to make convenient to generate a non-monotonic shape of response). This in fact shows that, with the right experimental conditions, it is possible to obtain a non-Piéron's shape with a Classic RT task, even though the "right experimental conditions" were not those predicted.

Different q conditions resulted in mdRT that were not statistically different from each other. However, in fitting the distributions, it appeared that the DDM using different a for different q conditions was the preferred model. This hinted to the fact that the effect of q was very weak, and only appeared in the response distribution, not in the mdRT. Even so, the difference across estimated a was really low, and not always increasing with increasing q. The most likely cause was that participants did not really put much weight on the points earned or lost, as the exchange rate between points and money was not considered favourable. Therefore, participants possibly used a unique (subjective) punishment/reward value, which was not considerably affected by the point system used. A replication with a more motivational exchange rate might test this possibility.

Finally, the optimality analysis in the context of DDM was clearly unsatisfying, for both aggregated and individual participants, providing with DT curve across accuracy level that did not resemble the inverted U shape expected in this case. In the other two works in which the optimal DT were computed (Bogacz et al., 2010; and Balci et al., 2011) the aggregated results did not show, similarly to our results,

a clear pattern. In Bogacz et al., a satisfying inverse U-shape was obtained for the top 30% and middle 60% performance; in Balci et al., when the data from the most extreme conditions were excluded, and only for the most advanced sessions (that is, after more than half an hour of training), the pattern of DT could be fit satisfactorily. It is possible that our sessions were too short to obtain a clear pattern, and participants were still varying their response in order to adjust the speed-accuracy trade-off. However, the time series analysis performed above did not result in any relevant pattern, which implied that participants were not slowly adjusting their response (inspection of individual responses also confirmed that). Another possible problem may reside in the aggregate analysis across q conditions, which may have mixed results from different strategies. The plot in Figure 8.12 shows that the difference across q conditions was (on average) very limited, so this possibility also seems unlikely. It was not possible to conduct an analysis on the results segregated by q conditions, as this resulted in too few data points, often having no entry for some level of accuracy. Another possibility is that, by fitting aggregating distributions, we missed relevant differences amongst participant's non-decision time, τ, which therefore resulted in badly scaled DT. However, our approach is not new (e.g. Simen et al., 2009), and it is preferred when dealing with several participants which produced less than one hundredo of trials for condition. Furthermore, by using a flexible $\widehat{D}_{tot}$ estimation, we should have been able to take into account slightly misestimation of τ. One of the main points of the present work was to suggest the use of an experimental design, the EXACT Paradigm, in which participants' strategy can be analysed regardless of the perceptual model used. The unsatisfying results in this study, coupled with the mixed results in the literature, strengthen this point.

Chapter 9

Summary and Conclusion

9.1 Summary of findings

The study of optimality has exerted a crucial impact on research, affecting significantly in the last decades the study of human behaviour, especially through economic and ecology. In this work, we approached the problem of optimality by trying to understand what optimal strategy is used by human participants in simple, fast decisions, mostly revolving around two alternatives. We limited our analysis to 4 strategies and, during the course of our exploration, we investigated several topics related to the wider field of decision making.

The first part of this work has been devoted to a single, but revealing, aspect of optimisation: that is the prediction that participants change their response based on the total duration of the condition. This non-intuitive prediction was actually supported by the previous literature (Bogacz et al., 2010; Balci et al., 2011; Simen et al., 2009).

We started by using a qualitative approach for Single and Choice RT task, showing how important is to employ the correct experimental paradigm. In particular, whereas the Single RT task in Chapter 4 appeared to support previous findings, the experiment in Chapter 5 provided a convincing rebuttal of this hypothesis, and also clarified the reason for alternative results in the literature. We realised that the previous experiments, by changing the RSI time, were affecting the stimulus variability and its connection to motor preparation, and thus the effect on participants was not the result of an optimisation strategy based on the RSI.

In Chapter 6, we tackled the problem of investigating several decision rules by basing the analysis on some perceptual model, which could be limiting or misleading. We suggested a new experimental

design in which participants were asked to trade between speed and accuracy in a similar way they have to do during a Classic RT task, but by using a speed-accuracy function that is perfectly known by the researcher. We found in this Chapter that, whereas this approach appeared powerful, the lack of information about the real $ACC(RT)$ made the response extremely noisy, which seriously affected the analysis. Therefore in Chapter 7 we employed a new version of this paradigm by showing the $ACC(RT)$ directly to the participants. We tested again, with this new paradigm, the hypothesis of delay effect on responses, confirming previous Chapters' results. We also suggest a different explanation, that is: participants started with an estimated D_{tot} which is the same across different delay conditions, and they changeed it only if this has a relevant effect in terms of profit. It is possible that, for our experiments, participant did not deem necessary to change their D_{tot} estimation, as this would not have resulted in a much higher utility. We did not explore this but strongly suggest it as a topic in a future study.

The next two studies (the last in Chapter 7 and the first in Chapter 8) we used the EXACT Paradigm to explore participants' strategy across several experimental conditions. These two studies, together with the last, comparative, Classic RT study, served two main purposes: analysing the decision rules used by participants and how they can be affected by experimental parameters, and investigate the effect in terms of Piéron's law.

9.2 Decision Rules and Optimality

The findings for the decision strategies were not easy to interpret. We generally found that the RR_m and the RA decision rules provided the best fit, with a slight preference for RR_m. It seems clear that the BR models could not account for the data; on the other hand, the RR model is a simplified version of RR_m, and it could fit the data only when the subjective punishment was low enough, but not in any general situation. Both RR_m and RA were been proposed by Bogacz et al. (2010) as a way to weight speed and accuracy. RR_m has an easy interpretation, as it is the strategy that maximizes the profit in terms of subjective utility given a task with fixed experimental time, but the RA does not, as it is an arbitrary function that weight differently accuracy and reward rate (Bogacz et al., 2006). Incidentally, both strategies are function of D_{tot} , which is peculiar, since we have clearly observed that participants

did not change their behaviour based on the trial length alone. This hints to the alternative suggested above: participants may vary their response depending on D_{tot}, but they vary it only when it is economically convenient, otherwise they use a default value. Another option, suggested in **Section 7.5**, could be that the parameter D_{tot}, used in both strategies, may have a different interpretation that the one usually suggested, and may thus not be related to total trial time. Finally, it is possible that participants are indeed using RR_m to maximize the reward rate, but that they are only considering a limited part of the trial (that is, they are using a different "epoch"), in which D_{tot} is computed. However, these two last options both imply that RR_m would lose its simple meaning as a rule that maximizes the reward rate for the trial. Another problem is that discrimination between RR_m and RA may not be simple, as these decision strategies predict similar responses for similar parameters. By comparing our results with the previous work (see **Section 8.1**) we realize that a clear difference between these two decision strategies cannot be easily drawn.

Why understanding the rules employed by human participants is so complex? We identified 3 main points:

1) Participants may change the parameters of their decision rules during the task itself. We have had several hints of this process. Participants may start with a default estimate of the parameters, and adjust them slowly along the session. We observed trends which seemed to suggest an adjustment towards a particular set of parameters (for example, see Figure 8.3). Furthermore, it has been observed, by using a design with several sessions (Balci et al., 2011) that at least one parameter, q, appeared to decrease with practice. Participants may change their parameter due to better estimation (which seems the more likely possibility for λ) or modification of internal state (for example, by becoming less averse to punishment, and thus decreasing q).

2) Different participants may use different parameters/decision rules. By aggregating participants, we basically estimated the parameters for an average participant, which may be misleading about the true participants' parameters. We tried to take this problem into account by separating participants into

performance group, but an individual analysis was not possible due to the small sample size for some conditions. This problem became even more serious when considering that different participants may use totally different strategies. For future studies, we suggest to conduct experiments with much larger sample size for each participant.

3) Over-estimation of Optimum Response Time (Bogacz et al., 2006). The characteristic asymmetric decision strategy curve may lead to many more losses when the optimum response is underestimated than overestimated, and participants are thus biased towards overestimation. This would produce RT that are higher than the optimum response RT^* and, thus, have higher accuracy. This process would cause misleading parameters fitting, for example resulting in a higher q that the one estimated by the participant. Taking this aspect into account involves calculating the optimum response giving the participant's noise estimation (Balci et al., 2011). Note that the same effect can be obtained by simply assuming that participants' response is intrinsically noisy: in this case as well, participants' best strategy would be to slightly increase the mean of their noisy response. The details of this hypothesis are highly speculative, and clearly more research about the threshold adaptation has to be done before this aspect can be fruitfully included in future analysis (see also Simen, Cohen & Holmes, 2006). To give an example of how this would affect optimum response estimation, in Figure 9.1 we assumed that the RT^* estimation is normally distributed[*] and, with different types of noise (σ), we calculated the optimum response time by using the decision rule RR_m. With $\sigma = 0$ the RT^* is exactly the maximum of the decision rule, as we have assumed in this work. With higher σ the optimum shifts on the right. For now, we noticed that ignoring these facets may lead to slightly incorrect parameter estimation.

[*] Note that, even if the RT^* estimation is based on a Normal Distribution process, the resulting RT distribution may still have the classic right-skewed shape, depending on the method used to adjust the response to find the optimum RT^*.

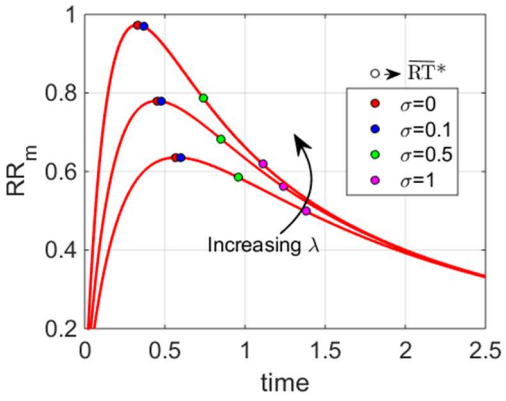


Figure 9.1

Optimum average response ($\overline{RT}^*$) for different values of response noise (coloured circles). Note that, as the noise increases, the optimum average response will shift on the left, meaning that participants need to be slower, and their average reward is going to be smaller.

For EXACT Paradigm, these represent the main problems. For the Classic RT task, we have a further problem of choosing the correct perceptual model (see Chapter 6 for discussion). Another complication comes from the fact that it has been shown that, for the Extended DDM, a model with linear boundary may actually be not optimal for maximising several decision strategies (RR and RR_m included), and a moving boundary must be used instead (Drugowitsch et al. 2012). Including a moving boundary adds a high level of freedom to the model, and makes it remarkably more complex. For our qualitative analysis in Chapter 4 and 5, we deemed a linear DDM a sufficient descriptive model, but clearly for any quantitative analysis this does not suffice. The analysis of optimality clearly depends on the development of this new type of sequential sampling models.

Finally, we must consider the possibility that other strategies may be used by participants. We have considered four decision rules that have been suggested several times across the literature. However, it is possible that participants may be using more complex strategies not taken into account here. For example, participants may base their response on a non-linear cost of waiting (Drugowitsch, 2012); the effect of the punishment/reward ratio may be modelled in a much complex way, especially by considering that participants may give different value to different punishment/reward *absolute* value, conversely to what assumed here.

9.3 Piéron's Law

In the last two Chapters, we designed the experiments such that one of the factors was always λ, which can be seen as corresponding to stimulus strength in a Classic RT task. This allowed us to perform studies that resembled Piéron's-type experiment. One point of this approach was to show the similarity between the EXACT Paradigm and the Classic RT task. However, the most interesting peculiarity of this approach was to obtain what we called a non-Piéron's shape with some experimental conditions, a shape that has not been observed in a century of RT experiments. Both experiments are clearly successful in generating both Piéron's and non- Piéron's shape by changing the parameter a and q (starting point and punishment/reward ratio). The Classic RT task also generated non-Piéron's shape, but only for the participants in the poor performance group. We suggested that participants from this group may actually have a poorest vision which would result in a weaker subjective stimulus strength which would produce the observed non-Piéron's shape.

Showing that Piéron's Law does not hold for certain experimental conditions has important psychological implications. This relationship between stimulus strength and response time has been so ubiquitous that is normally considered one of the main psychophysical regularities, and it has seen as the result of the neural activation of the visual system. For example, Stafford and Gurney (2004) and Stafford, Ingram and Gurney (2011) suggested that Piéron's Law may be generated by a sequential sampling model (they used a Linear Ballistic Model, but the same prediction holds with all the major sequential sampling models). Another approach is the Ising Decision Maker, derived from a stochastic Hopfield network (Verdonck and Tuerlinckx, 2014). However, such models cannot, by themselves, explain the non-monotonic relationship that we have observed in Experiment 7.2 and 8.1, and 8.2.

By showing that Piéron's Law can easily be broken, and specifying when this happens, we showed how even this phenomena, considered purely physiological, are in fact mediated by participants' decision rules. We are not claiming that visual activation does not have a role here: in fact, the stimulus strength clearly affects the speed-accuracy function, which is employed by participants to estimate the

decision rule value. What we claim is that the observed response can only be deeply understood by taking into account the decision rule in which the physiological response is applied.

9.4 Rate-Hypothesis and Accuracy-Hypothesis

Finally, during this work, we have investigated two important phenomena: the rate of response and the accuracy of response. The first regularity, that is that the distributions of response rate are approximately Normal, arose in the literature of saccadic response and it has been confirmed for Manual Choice RT (see **Section 2.4**). As for the accuracy distribution, we noticed for the first time in Experiment 7.1, only within the context of the EXACT Paradigm, that it was also Normally distributed. The fact that a regularity in the accuracy distribution may have been overseen by previous researcher is not surprising, as accuracy distributions may be difficult to compute.

For the Rate Domain Hypothesis, first suggested by Harris et al. (2014), we refer to the hypothesis that the decision making process is taking place in the Rate Domain instead of the time domain. The DDM and other sequential sampling processes would be interpreted as a way to convert from rate to time domain. The LATER process is the "pure" version of this assumption, as it starts from a Normal Distribution in the rate domain and generates a Reci-Normal Distribution in in time domain. We have shown how this hypothesis is very difficult to prove as it is generally difficult to distinguish between a distribution in the rate domain generated by a DDM and a pure Normal Distribution in the rate domain (see **Section 5.4**). Nevertheless, we observed for both Simple RT and Choice RT (Chapter 4 and 5) that the responses in the rate domain were in fact approximately Normally distributed, but with some consistent deviations that may be due to anticipatory processes/stimulus estimation. The deviations were present specifically when the FP was short (or constant), hinting that participants were sometimes basing their estimation on the FP itself instead that on the perceptual process. The LATER model, on the other hand, was compared with the DDM. The latter appeared to provide a better distribution fit. Overall, we believe that the information gathered in this work did not provide strong support to the Rate-Hypothesis, and more work is clearly required.

An important result that we believe deserve to be explored more consistently is the regularity found by analysing the distribution accuracy. We realised that participants may indeed compute their response based on a particular value of accuracy that they wish to maintain, and the produced RT is only a sub-product of such accuracy. We showed by simulations (**Section 7.10**) how this could explain several distribution features. This is an extremely simple model which takes distance from a sequential sampling type of model and obviously requires much more investigation. This regularity has been consistently found across our EXACT experiments, but an analysis on Classic RT is yet to be performed. This analysis can be very complicated, and requires probably thousands of trials, but could be possible with an experiment designed just to extrapolate this information. Finding normally distributed accuracy distributions would be such a peculiar coincidence that would most likely hint to some more important mechanism underlying the decision process. Note that, by introducing this assumption, we are not suggesting any radical change to the optimisation process adopted until this moment: participants may compute the optimum response (in time) given a decision strategy. They will then respond based on the accuracy around that point, and this accuracy may be in fact normally distributed. Note that the RT probability density function suggested in **Section 7.10** (Equation 7.1) is in fact function of the optimum response time RT^* and the accuracy related to it, $ACC(RT^*)$. Noisy estimation of the accuracy, or motor noise, or any other normally distributed process, would then take place on the accuracy domain.

Why is this assumption so relevant? The main point regards to the fact that, as discussed in **Section 6.1.1**, perceptual models and decision rules are two separate components, and they can, in fact, be analysed separately. This is exactly what we did by employing the EXACT Paradigm. Once we conceptually accept this separation, we can then consider alternative models for generating the RT distributions, and they may not be sequential sampling models. In this case, assuming that participants were responding so to maintain a certain level of accuracy would automatically imply that a right skewed distribution will be generated in the time domain. This is a new, simple explanation for RT

distribution which does not employ any sequential sampling model, but only a certain $ACC(RT)$ shape and a decision rule.

Part of this thesis has to be considered exploratory: as we proposed a new paradigm, it opens up several possibilities for further research with an in-depth investigation of few of them. Understanding how humans make decisions is a difficult endeavour, even when we limit our analysis to simple and fast decisions as the one performed in reaction time tasks. It sometimes appears necessary to employ novel experimental designs that can help us shed some light onto the intricate landscape of decision models and decision strategies. The key point of this work was to show the limit of the current approach with Classic RT task and to suggest a possible solution in terms of the EXACT Paradigm. Exploration of this novel design gave us new information about the decision strategies employed in human perception, but this will clearly require further investigation. By suggesting new directions for increasing our knowledge of decision making, we hope to have provided a relevant contribution to the field.

Appendix

Derivation of Optimum Response DT^* for the four decision rules

For each decision rule, we describe the steps to derive the corresponding t^* for the exponential speed-accuracy trade-off function: $ACC(DT) = 1 + (\alpha - 1)\exp(-\lambda DT)$. To find the maximum point in each decision rule, we set the derivative with respect to t equal to 0 and solve for t.

For $BR = -\left[DT + q\left(1 - ACC(DT)\right)\right]$, the derivative is

$$\frac{\partial BR}{\partial DT} = -1 - \lambda q\, exp(-\lambda DT)(\alpha - 1) = 0 \tag{A1}$$

and hence:

$$DT^*_{BR} = -\frac{ln\left(\frac{1}{\lambda q(1-\alpha)}\right)}{\lambda}. \tag{A2}$$

Applying L'Hopital's rule, we found that when $\lambda \to 0$ then $DT^*_{BR} \to 0$ and when $\lambda \to \infty$ then $DT^*_{BR} \to 0$.

For $RR = ACC(DT)/(DT + d)$ the derivative is

$$\frac{\partial RR}{\partial DT} = -\frac{1 - exp(-\lambda DT)\,(\alpha - 1)}{(DT + D_{tot})^2} - \frac{\lambda\, exp(-\lambda DT)\,(\alpha - 1)}{DT + D_{tot}} \tag{A3}$$

$$= 0$$

Rearranging:

$$(1 + \lambda t + \lambda D_{tot})\, exp(-\lambda DT) = -\frac{1}{\alpha - 1} \tag{A4}$$

and setting :

$$x = -\lambda DT - \lambda D_{tot} - 1 \tag{A5}$$

we obtain $x\, exp(x + \lambda D_{tot} + 1) = \frac{1}{\alpha - 1}$ or

$$x\, exp(x) = \frac{exp(-\lambda D_{tot} - 1)}{\alpha - 1} \tag{A6}$$

We can now express this relationship in terms of Lambert W function, which is defined as $W(z)\exp[W(z)] = z$. Therefore

$$x = W\left(exp\frac{(-\lambda D_{tot} - 1)}{\alpha - 1}\right) \tag{A7}$$

which we can substitute back to (A5) to obtain

$$DT^*_{RR} = -D_{tot} - \frac{W\left(\frac{exp(-\lambda D_{tot} - 1)}{\alpha - 1}\right) + 1}{\lambda} \tag{A8}$$

As the Lambert W function is a multivalued function, we need to establish which branch we are referring to. Note that to obtain $DT^*_{RR} \geq 0$ we need $W(\cdot) \leq -1 - \lambda D_{tot}$. Since $\lambda, D_{tot} \geq 0$, then $W(\cdot) \leq -1$ and therefore the function is limited to its lower branch, W_{-1}. With $\lambda \to 0$ RR become $\frac{\lambda}{DT + D_{tot}}$ and the maximum is reached when $DT = 0$.

We were not able to obtain an explicit form for RA, not even in terms of Lambert W.

For $RR_m = \frac{ACC(DT) - q(1 - ACC(DT))}{DT + D_{tot}}$ the derivative is equal to

$$\frac{\partial RR_m}{\partial DT} = -\frac{exp(-\lambda DT)\,(\alpha-1) + q\,exp(-\lambda DT)\,(\alpha-1) + 1}{(DT + D_{tot})^2} \qquad \text{(A9)}$$

$$-\frac{\lambda\,exp(-\lambda DT)\,(\alpha-1) + \lambda q\,exp(-\lambda DT)\,(\alpha-1)}{DT + D_{tot}}$$

Which maximum DT can be found by setting this derivative equal to zero. By rearranging terms, we obtain:

$$(\lambda DT + \lambda d + 1)\,exp(-\lambda DT) = -\frac{1}{(\alpha-1)(q+1)} \qquad \text{(A10)}$$

As before, we set $x = -\lambda DT - \lambda d - 1$, so that we can express x in terms of Lambert W function:

$$x = W\left(\frac{exp(-\lambda D_{tot} - 1)}{(\alpha-1)(q+1)}\right) \qquad \text{(A11)}$$

Leading to

$$DT^*_{RR_m} = -D_{tot} - \frac{W\left(\frac{exp(-\lambda D_{tot} - 1)}{(\alpha-1)(q+1)}\right) + 1}{\lambda} \qquad \text{(A12)}$$

For the same argument as before, we know that for $DT^*_{RR_m}$ to be positive the function W is limited to its lower branch, W_{-1}. When $\lambda \to 0$ RR_m become $\frac{\alpha(1+q)-q}{DT+D_{tot}}$ which maximum depends on the sign of the numerator. It is $DT^*_{RR_m} = 0$ when $\alpha > \frac{q}{q+1}$, $DT^*_{RR_m} = \infty$ when $\alpha < \frac{q}{q+1}$, undetermined when $\alpha = \frac{q}{q+1}$.

References

Aiken, L. R., & Lichtenstein, M. (1964). Reaction times to regularly recurring visual stimuli. Percept. Mot. Skills, 18, 713-720.

Audley R. J. (1957). A stochastic description of the learning behaviour of an individual subject. Q. J. Exp. Psychol. 9, 12–20 10.1080/17470215708416215

Audley R. J. (1958). The inclusion of response times within a stochastic description of the learning behavior of individual subjects. Psychometrika 23, 25–31 10.1007/BF02288976

Balci, F., Simen, P., Niyogi, R., Saxe, A., Holmes, P., & Cohen, J. D. (2011). Acquisition of decision making criteria: Reward rate ultimately beats accuracy. Attention, Perception and Psychophysics , 73 (2), 640 – 657.

Banks W. (1973). Reaction time as a measure of summation of warmth. Percept. Psychophys. 13, 321–32710.3758/BF03214147

Bausenhart, K. M., Rolke, B., Seibold, V. C., & Ulrich, R. (2010). Temporal preparation influences the dynamics of information processing: Evidence for early onset of information accumulation. Vision Research, 50, 1024-1035.

Bernard, G. A. (1946). Sequential tests in industrial statistics, Journal of the Royal Statistical Society, 8 (1), Suppl. 1–21.

Bevan, W., Hardesty, D., & Avant, L. (1965). Response latency with constant and variable interval schedules. Perceptual & Motor Skills. 20, 969-972.

Blank, H., Biele, G., Heekeren, H. R., Philiastides, M. G. (2013). Temporal Characteristics of the Influence of Punishment on Perceptual Decision Making in the Human Brain. The Journal of Neuroscience, February 27, 2013. 33(9):3939 –3952

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychological review, 113(4) 700-765. doi:10.1037/0033-295X.113.4.700

Bogacz, R., Hu, P. T., Holmes, P. J., Cohen, J. D. (2010). Do humans produce the speed-accuracy trade-off that maximizes reward rate?. Quarterly journal of experimental psychology (2006), 63(5) 863-891. doi:10.1080/17470210903091643

Bonnet, C., Zamora, M. C., Buratti, F., Guirao, M. (1999). Group and individual gustatory reaction times and Pieron's Law. Physiology and Behavior, 66(4) 549-558. doi:10.1016/S0031-9384(98)00209-1

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. Psychological Review, 85, 1-21.

Brown, S. D., Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. Cognitive Psychology, 57(3) 153-178. doi:10.1016/j.cogpsych.2007.12.002

Briggs, G. E., & Blaha, J. (1969). Memory retrieval and central comparison times in information processing. Journal of Experimental Psychology, 79, 395-402

Carpenter, R. H. S. (1981). “Oculomotor procrastination,” in Eye Movements: Cognition and Visual Perception, eds D. F. Fisher, R. A. Monty, and J. W. Senders (Hillsdale: Lawrence Erlbaum), 237–246.

Carpenter, R. H. S. (2004). Contrast, probability, and saccadic latency: evidence for independence of detection and decision. Curr. Biol. 14:1576–80

Carpenter R. H., Reddi B.A., Anderson A.J. (2009).A simple two-stage model predicts response time distributions. J. Physiol. 587(Pt 16):4051-62

Carpenter, R. H., Williams, M. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. Nature, 377(6544) 59-62. doi:10.1038/377059a0

Donders F. C. (1868). On the speed of mental processes. Arch. Néerland. 3, 269–317

Donkin C., Brown S., Heathcote A., Wagenmakers E. (2011). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? Psychon. Bull. Rev. 18, 61–6910.3758/s13423-010-0022-4

Donkin, C., & Van Maanen, L. (2014). Piéron's Law is not just an artifact of the response mechanism. Journal of Mathematical Psychology, 62-63, 22-32.

Diederich, A., (2008). A further test of sequential-sampling models that account for payoff effects on response bias in perceptual decision tasks. Perception & Psychophysics, 70, 229-256, 2008 http://dx.doi.org/10.3758/PP.70.2.229

Diederich A., & Busemeyer J.R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time and simple response time. Journal of Mathematical Psychology;47:304–322

Drazin, D. (1961). Effect of foreperiod, foreperiod variability, and probability of stimulus occurrence on simple reaction time. Journal of Experimental Psychology. 62, 43-50.

Drugowitsch J., Moreno-Bote R., Churchland A.K., Shadlen M.N., Pouget A. (2012). The cost of accumulating evidence in perceptual decision making. J. Neurosci. 32:3612–28

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. Psychonomic Bulletin & Review, 16 , 1026–1036

Edwards W. (1954). The theory of decision making. Psychol. Bull. 51, 380–417 10.1037/h0053870

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. Psychological Review, 70, 193-242.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: Extremely short reaction times of goal directed eye movements. Experimental Brain Research. 57. doi:10.1007/BF00231145

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. Annual Review of Psychology, 67, 641-666.

Garrett H. E. (1922). A study of the relation of accuracy to speed. Arch. Psychol. 56, 1–106

Gold, J. I., Shadlen, M. N. (2002). Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. Neuron, 36(2) 299-308. doi:10.1016/S0896-6273(02)00971-6

Goldfarb S., Leonard N. E., Simen P., Caicedo-Nez C. H., Holmes P. (2014). A comparative study of drift diffusion and linear ballistic accumulator models in a reward maximization perceptual choice task. Front. Neurosci. 8:148. 10.3389/fnins.2014.00148

Granjon, M., Requin, J., Durup, H., & Reynard, G. (1973). Effects of timing signal on reaction time with "non-aging" foreperiods. Journal of Experimental Psychology. 101, 139-145.

Green, D. M. & Luce, R. D. (1974). Timing and counting mechanisms in auditory discrimination and reaction time. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (eds.), Contemporary Developments in Mathematical Psychology, Vol. II. San Francisco: Freeman, pp. 372-415.

Green, D. M., A. F. Smith, & S. M. von Gierke (1983). Choice reaction time with a random foreperiod. Perception & Psychophysics, 34, 195-208.

Grosjean, M., Rosenbaum, D. A., and Elsinger, C. (2001). Timing and reaction time. J. Exp. Psychol. Gen. 130, 256–272. doi: 10.1037/0096-3445.130.2.256

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. Science, 274, 427– 430.

Harris, C. M., and Waddington, J. (2012). On the convergence of time interval moments: caveat sciscitator. J. Neurosci. Methods 205, 345–356. doi: 10.1016/j.jneumeth.2012.01.017

Harris, C.M., Waddington, J., Biscione, V., Manzi, S. (2014). Manual choice reaction times in the rate-domain. Frontiers in human neuroscience, (June). doi:10.3389/fnhum.2014.00418

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions.Psychonomic Bulletin & Review,9, 394–401.

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. Frontiers in Neuroscience (June). doi:10.3389/fnins.2014.00150

Henmon V. A. C. (1911). The relation of the time of a judgment to its accuracy. Psychol. Rev. 18, 186–201 10.1037/h0074579

Hermelin, B. (1964). Effects of variation in the warning signal on reaction times of several subnormals. Quarterly Journal of Experimental Psychology. 16, 241-249.

Hick, W. E. (1952). On the rate of gain of information. Quarterly Journal of Experimental Psychology. 4 (1): 11–26. doi:10.1080/17470215208416600.

Holland, P. C., & Gory, J. (1986). Extinction of inhibition after serial and simultaneous feature negative discrimination training. Quarterly Journal of Experimental Psychology, 38K, 245-265.

Holmes, P., Cohen, J. D. (2014). Optimality and some of its discontents: Successes and shortcomings of existing models for binary decisions. Topics in Cognitive Science, 6(2) 258-278. doi:10.1111/tops.12084

Jaskowski, P., & Sobieralska, K. (2004). Effect of stimulus intensity on manual and saccadic reaction time. Percept. Psychophys. 66, 535-544.

Jepma, M., Wagenmakers, E.-J., & Nieuwenhuis, S. (2012). Temporal expectation and information processing: A model-based analysis. Cognition. 122,426-441.

Jones P.H., Harris C.M., Woodhouse J.M., Margrain T.H., Ennis F.A., Erichsen J.T. (2013) Stress and visual function in infantile nystagmus syndrome. Invest. Ophthalmol. Vis. Sci. 54(13):7943-51. doi: 10.1167/iovs.13-12560.

Karlin, L. (1959). Reaction time as a function of foreperiod duration and variability. Journal of experimental psychology. 58 185-191. doi:10.1037/h0049152

Klemmer, E. T. (1956). Time uncertainty in simple reaction time. Journal of Experimental Psychology. 51, 179-184.

Klemmer, E. T. (1957). Simple reaction time as a function of time uncertainty. Journal of Experimental Psychology. 54, 195-200.

Krajbich, I., Armel, K. C., and Rangel, A. (2010).Visual fixations and the computation and comparison of value in simple choice. Nat. Neurosci. 13, 1292–1298

LaBerge, D. (1962). A recruitment theory of simple behaviour. Psychometrika. 27(4) 375-396

LaBerge D. (1994) Quantitative models of attention and response processes in shape identification tasks. Journal of Mathematical Psychology.38:198–243.

Laming, D. R. J. (1968). Information theory of choice-reaction times. New York: Wiley.

Leite FP. (2012) A comparison of two diffusion process models in accounting for payoff and stimulus frequency manipulations. Atten. Percept. Psychophys. 74: 1366–1382

Link S.W. (1975). The relative judgement theory of two choice response time. Journal of Mathematical Psychology. 12:114–135.

Link S.W., & Heath R.A. (1975). A sequential theory of psychological discrimination. Psychometrika. 40:77–105.

Liston D. B. Stone L. S. (2008). Effects of prior information and reward on oculomotor and perceptual choices. Journal of Neuroscience, 28 (51), 13866–13875

Luce, R. (1986). Response Times: Their Role in Inferring Elementary Mental Organization. Oxford University Press New York. doi:10.1093/acprof:oso/9780195070019.001.0001

Maanen, L., Grasman, R. P. P. P., Forstmann, B. U., Keuken, M. C., Brown, S. D., Wagenmakers, E. (2012). Similarity and number of alternatives in the random-dot motion paradigm. Attention, Perception, & Psychophysics, 74(4) 739-753. doi:10.3758/s13414-011-0267-7

Maddox, W. T., & Bohil, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 1459-1482.

Mansfield, R. J. W. (1973). Latency functions in human vision. Vision Research. 13, 2219-2234.

Martin L. J., Müeller G. E. (1899). Zur Analyse der Unterschiedsempfindlichkeit. Leipzig: J. A. Barth.

Medina J. M. (2009). 1/f (alpha) noise in reaction times: a proposed model based on Pieron's law and information processing. Phys. Rev. E 79:011902 10.1103/PhysRevE.79.011902.

Merkel, J. (1885). Die Zeitlichen Verhältnisse der Wiellensthätigkeit. PhilosophischeStudien,2, 73–127.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., and Rangel, A. (2010).The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. Judgm. Decis. Mak.5, 437–449.

Moran, O. (2014). Optimal decision making in heterogeneous and biased environments. Psychonomic Bulletin & Review. 06/2014; 22(1). DOI: 10.3758/s13423-014-0669-3.

Mulder MJ, Bos D, Weusten JMH, van Belle J, van Dijk SC, et al. (2010). Basic impairments in regulating the speed-accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. Biol. Psychiatry 68:1114–19.

Nachmias J., Kocher E. C. (1970). Visual detection and discrimination of luminance increments. J. Opt. Soc. Am. 60, 382–389.

Näätänen, R. (1970). The diminishing time-uncertainty with the lapse of time after the warning-signal in reaction-time experiments with varying fore-periods. Acta Psychologica. 34, 399-419.

Näätänen, R. (1971). Non-aging fore-period and simple reaction-time. Acta Psychologica. 35, 316-327.

Nickerson, R., & Burnham, D. (1967). Response times with nonaging foreperiod. Journal of Experimental Psychology. 79, 452-457.

Nickerson, R. S. (1967). Psychological refractory phase and the functional significance of signals. Journal of Experimental Psychology. 73, 303-312.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. Psychol. Bull. 89, 133–162. doi: 10.1037/0033-2909.89.1.133.

Noorani, I. (2014). LATER models of neural decision behaviour in choice tasks. Front. Integr. Neurosci. 8:67. doi:10.3389/fnint.2014.00067.

Noorani I., Carpenter R. H. S. (2011). Full reaction time distributions reveal the complexity of neural decision-making. The European Journal of Neuroscience, 33 (11), 1948–1951 doi:10.1111/j.1460-9568.2011.07727.x.

Overbosch, P., Wijk, R. d., Jonge, T. J. d., Köster, E. P. (1989). Temporal integration and reaction times in human smell. Physiology & behavior, 45(3) 615-626. doi:10.1016/0031-9384(89)90082-6.

Patanaik A., Zagorodnov V., Kwoh C.K., & Chee M.W. (2014). Predicting vulnerability to sleep deprivation using diffusion model parameters. J. Sleep Res. 2014 Oct;23(5):576-84. doi: 10.1111/jsr.12166.

Piéron, H. (1914). Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices [Evidence on the laws of sensory processing time as a function of the excitatory intensity]. L'Année Psychologique. 20, 17–96.

Piéron, H. (1920). Nouvelles recherches sur l'analyse du temps de latence sensorielle en fonction des intensités excitatrices [Further evidence on the laws of sensory processing time as a function of the excitatory intensity].L'Année Psychologique. 22, 58–142.

Piéron, H. (1952). The sensations: Their functions, processes and mechanisms. London: Frederick Muller Ltd.

Pike R. (1966). Stochastic models of choice behaviour: Response probabilities and latencies of finite Markov chain systems. British Journal of Mathematical and Statistical Psychology. 21:161–182.

Pike R. (1973) Response latency models for signal detection. Psychological Review. 80:53–68.

Pins, D., Bonnet, C. (1996). On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time. Percept. Psychophys, 58(3) 390-400. doi:10.3758/BF03206815.

Pins, D., Bonnet, C. (1997). Reaction times reveal the contribution of the different receptor components in luminance perception. Psychonomic Bulletin & Review, 4(3) 359-366. doi:10.3758/BF03210793.

Pins, D., & Bonnet, C. (2000). The Piéron function in the threshold region. Perception & psychophysics, 62(1), 127-136.

Ratcliff, R. (1978). A theory of memory retrieval. Psychological Review, 85(2) 59-108. doi:10.1037/0033-295X.85.2.59.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. Psychological bulletin, 86(3) 446-461. doi:10.1037/0033-2909.86.3.446.

Ratcliff R., & van Dongen H.P.A. (2011) A diffusion model for one-choice reaction time tasks and the cognitive effects of sleep deprivation. Proceedings of the National Academy of Sciences. 108:11285–11290.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. Psychological review, 111(1) 159-182. doi:10.1037/0033-295X.111.1.159

Ratcliff R., Hasegawa Y.T., Hasegawa Y.P., Smith P.L., Segraves M.A. (2007). Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. Journal of Neurophysiology. 97:1756–1774.

Ratcliff, R., Murdock, B. B. Jr. (1976). Retrieval processes in recognition memory. Psychological Review, 83, 190-214.

Ratcliff R, Perea M, Colangelo A, & Buchanan L. (2004) A diffusion model account of normal and impaired readers. Brain & Cognition. 55:374–382.

Ratcliff, R., & Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. Psychological Science, 9(5) 347-356. doi:10.1111/1467-9280.00067

Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. Journal of experimental psychology. Human perception and performance, 26(1) 127-140. doi:10.1037/0096-1523.26.1.127

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. Psychological review, 111(2) 333-367. doi:10.1037/0033-295X.111.2.333

Ratcliff, R., & Strayer, D. (2014). Modeling simple driving tasks with a one-boundary diffusion model. Psychonomic Bulletin and Review, 21, 577–589.

Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on brightness discrimination. Perception & Psychophysics, 65, 523–535

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating the parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. Psychonomic Bulletin and Review. 9:438–481.

Ratcliff R, Van Zandt T, McKoon G. Connectionist and diffusion models of reaction time. Psychological Review. 1999;106:261–300.

Reddi B.A., Asrress K.N., Carpenter R.H. (2003). Accuracy, information, and response time in a saccadic decision task. J. Neurophysiol. 90(5):3538-46.

Reddi, B. A. J., & Carpenter, R. H. S. (2000). The influence of urgency on decision time. Nature Neuroscience. 3, 827-830.

Rodionov, S. N., (2004). A sequential algorithm for testing climate regime shifts. Geophys. Res. Lett., 31, L09204, doi:10.1029/2004GL019448.

Rodionov, S. N., & Overland J. E. (2005). Application of a sequential regime shift detection method to the Bering Sea ecosystem. ICES J. Mar. Sci., 62: 328-332.

Rolke, B., & Hofmann, P. (2007). Temporal uncertainty degrades perceptual processing. Psychonomic Bulletin & Review, 14, 522 – 526. doi:10.3758/BF03194101.

Saslow, C. A. (1968) Operant control of response latency in monkeys: Evidence for a central explanation. Journal of the Experimental Analysis of Behavior, 11, 89-98.

Saslow, C. A. (1972) Behavioral definition of minimal reaction time in monkeys. Journal of the Experimental Analysis of Behavior, 18,87-106.

Schall, J. D., & Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. Annual Review of Neuroscience, 22, 241–259.

Seibold, V. C., Bausenhart, K. M., Rolke, B., & Ulrich, R. (2011). Does temporal preparation increase the rate of sensory information accumulation? Acta Psychologica, 137, 56–64.

Shadlen, M. N., Newsome, W. T. (2001). Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey. J. Neurophysiol. 86(4) 1916-1936. doi:10.1016/j.neuroimage.2011.08.034

Simen, P., Cohen, J. D. (2009). Explicit melioration by a neural diffusion model. Brain Research. 1299 95-117. doi:10.1016/j.brainres.2009.07.017.

Simen P., Cohen, J.D., Holmes P. (2006). Rapid decision threshold modulation by reward in a neural network. Neural Netw. Oct; 19(8): 1013-26.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., & Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: Empirical tests of theoretical predictions. Journal of Experimental Psychology: Human Perception and Performance. 35 (6), 1865–1897.

Snodgrass, J. G., Luce, R. D., & Galanter, E. (1967) Some experiments on simple and choice reaction time. Journal of Experimental Psychology. 75, 1-17.

Stafford, T., Gurney, K. N. (2004). The role of response mechanisms in determining reaction time performance: Piéron's law revisited. Psychonomic bulletin & review, 11(6) 975-987. doi:10.3758/BF03196729.

Stafford, T., Ingram, L., & Gurney, K. N. (2011). Piéron's Law holds during stroop conflict: Insights into the architecture of decision making. Cognitive Science. 35(8), 1553-1566.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. Acta Psychologica. 30: 276–315. doi:10.1016/0001-6918(69)90055-9.

Stone, M. (1960). Models for choice-reaction time. Psychometrika, 25(3) 251-260. doi:10.1007/BF02289729

Speckman, P. L., & Rouder, J. N. (2004). A comment on Heathcote, Brown, and Mewhort's QMLE method for response time distributions. Psychonomic Bulletin & Review, 11, 580-582.

Spencer, T. J., & Chase, P. N. (1996). Speed analysis of stimulus equivalence. Journal of the Experimental Analysis of Behavior. 65, 643-659.

Townsend, J.T., & Ashby, F.G. (1983). Stochastic modeling of elementary psychological processes. Cambridge: Cambridge University Press.

Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty. 5, 297–323.

Ulrich, R., & Miller, J. (1994). Effects of truncation of reaction time analysis. Journal of Experimental Psychology: General. 123 , 34-80.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. Psychological review. 108(3) 550-592. doi:10.1037/0033-295X.108.3.550.

Vallesi, A., Shallice, T., Walsh, V. (2006). Role of the prefrontal cortex in the foreperiod effect: TMS evidence for dual mechanisms in temporal preparation. Cereb Cortex. 2007 Feb; 17(2):466-74.

van Ravenzwaaij D, Brown S, & Wagenmakers EJ. (2011). An integrated perspective on the relation between response speed and intelligence. Cognition 119:381–93.

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. Ergonomics, 13, 37–58.

Vickers, D. (1978). An adaptive module of simple judgements. In J. Requin (Ed.), Attention and performance, VII (pp. 598– 618). Hillsdale, NJ: Erlbaum.

Vickers, D. (1979). Decision processes in visual perception. New York: Academic Press.

Vickers D., Caudrey D., & Willson R.J. (1971) Discriminating between the frequency of occurrence of two alternative events. Acta Psychologica. 35:151–172.

Vincent, S. B. (1912). The function of vibrissae in the behavior of the white rat. Behavioral Monographs , 1.

Verdonck, S., & Tuerlinckx, F. (2014). The Ising Decision Maker: A binary stochastic network for choice response time. Psychological Review, 2014[Jul], Vol 121[3], 422-462.

Wald, A. (1947). Sequential analysis. New York: Wiley.

Wald, A., & Wolfowitz, J. (1948). Optimum Character of the Sequential Probability Ratio Test. The Annals of Mathematical Statistics, 19(3) 326-339. doi:10.1214/aoms/1177730197.

Wagenmakers, E.J., & Brown S. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. Psychol. Rev. 114:830–41.

Wagenmakers, E.J., van der Maas, H.L., & Grasman, R.P. (2007). An EZ-diffusion model for response time and accuracy. Psychon. Bull. Rev. 14:3–2.

Wagenmakers, E.J., Ratcliff R., Gomez P., & McKoon G. (2008). A diffusion model account of criterion shifts in the lexical decision task. J. Mem. Lang. 58:140–59.

Whelan, R. (2008). Effective analysis of reaction time data. The Psychological Record. 58(3) 475-482.

Woodworth,R. S.(1899). The accuracy of voluntary movement. Psychological Review. 3, (3, Suppl. 13), 1-119.

Woodworth, R. S. & Schlosberg, H. (1954). Experimental Psychology. New York: Holt.

Zacksenhouse, M., Bogacz, R., & Holmes, P. (2010). Robust versus optimal strategies for two-alternative forced choice tasks. Journal of Mathematical Psychology. 54, 230–246.

Copies of publications

Decision Making and Rate of Reward

Valerio Biscione, Christopher M. Harris

Centre for Robotics & Neural Systems and Cognition Institute, Plymouth University, Plymouth, UK

ABSTRACT

It is widely accepted that human reaction times (RT) can reveal important information about the way we make decisions. It is commonly accepted that for simple decision a response is made when the accumulation of information reaches a threshold. In the present work we analyze the frequency distributions of the RT in order to understand how responses are made and, in particular, if responses are mediated by a decision mechanism whose purpose is to maximize the reward. We suggest that the important element that human beings are trying to maximize is not the reward per se, but the rate of reward, that is the reward per unit time. We conducted an experiment where we varied the intensity of the stimulus and the foreperiod time (which is the time between the start of the trial to the occurrence of the stimulus). We compared our result with other decision-making models and we showed how our experimental results are easy to understand within the framework of a reward-rate maximization mechanism and how this can provide new insight in the decision-making field for human beings.

For the *6th Annual Plymouth University School of Psychology Conference*. Plymouth. 2015.
(Oral Presentation)

Rate Domain Distributions in Simple Reaction Time Task

Valerio Biscione, Christopher M. Harris

Centre for Robotics & Neural Systems and Cognition Institute, Plymouth University, Plymouth, UK

ABSTRACT

Background. It is widely accepted that human reaction times (RT) can reveal important information about decision strategies for "all-or-none" type of responses. During the past decades several models have been proposed to describe and predict mean RT and RT distributions and, amongst them, sequential sampling models are the most successful in accounting for data for simple and two-choice tasks. RT distributions are usually positively skewed which are often modelled as an Inverse Gaussian (the first passage time of a stochastic Wiener process) or an Ex-Gaussian. However it has been found that latency distribution of saccades are very close to the Reciprocal Normal, meaning that they are normally distributed in the rate domain. This represents a challenge for most stochastic rise-to-threshold models, since these models can only achieve normality in the rate domain with implausible parameters. We have already observed, in a previous study, the normality of the distribution in the rate domain. In that experiment we used a two-forced choice paradigm with an easy/difficult condition and an accurate/urgent instruction sets. We now investigate the possibility to find similar results with a simple RT paradigm, exploring the relation between the foreperiod (FP) time and the Piéron's Law in the rate domain.

Methods. In this experiment (12 subjects, 3 blocks of 250 trials) we varied the FP time and the luminance of the stimuli. The participants were asked to press a button as soon as they saw the stimulus (a circle). We used 3 FP conditions (0.6, 1, and 2.4 seconds) and 5 luminance levels (0.42, 0.71, 1.21, 2.06, 3.50 cd/m^2).

Discussion. We discuss several problems of the rise-to-threshold models and how a Reciprocal Normal Distribution is not consistent with stochastic rise to threshold models. We propose a simple optimality model in which reward is maximized to yield to an optimal rate and therefore an optimal time to respond, exploring the connection between this model and the speed-accuracy trade-off. We also show how Piéron's Law naturally arises from this model. Our key claim is that the main goal of the human decision process in simple decision tasks is to maximize the rate of reward.

For the *First Decision Making Conference, 2014*. Bristol. *(Oral Presentation)*

How to Investigate decision-making in humans?

Valerio Biscione, Christopher M. Harris

Centre for Robotics & Neural Systems and Cognition Institute, Plymouth University, Plymouth, UK

ABSTRACT

Whether and in which conditions humans' decision making is optimal in reaction time tasks has been a long discussed question in the psychological field. The topic is affected by a vague definition of optimality, and by a lack of experimental paradigm designed to investigate the problem. Most of the models proposed in the past decades have focused their attention on perceptual tasks, where participants are asked to choose as fast as possible between two or more stimuli. These studies mostly ignored what strategy was used by the participants. However, we believe that a different approach could be employed: we introduce a novel paradigm in which participants are asked to make a decision and gain a reward by responding to a task which does not contain any perceptual cue. This allows us to analyse the strategy employed without making any assumption about the perceptual mechanism. We discuss how this paradigm could improve our knowledge about the decision strategy used by humans, and we present some preliminary results that show how the participants were trying to optimize the rate of reward from trial to trial.

For the *Plymouth Postgraduate Conference Series, 2015.* Plymouth. *(Oral Presentation)*

CONNECTING PIÉRON'S LAW, THE FOREPERIOD EFFECT AND DISTRIBUTION SHAPES IN A SIMPLE REACTION TIME TASK

Valerio Biscione[1,2], Christopher M. Harris[1,2]
[1]*School of Psychology, Plymouth University, Plymouth, UK*
[2]*Centre for Robotics and Neural Systems and Cognition Institute, Plymouth University, Plymouth, UK*

Abstract

In a recent study on choice reaction time (RT) (Harris et al., 2014), we found that rate (reciprocal of RT) was near-normally distributed, which led us to propose a decision model based on maximizing the reward rate (the Rate Model). We also suggested a connection with two well-known phenomena in experimental psychology: Piéron's Law and the foreperiod (FP) effect. Piéron's Law describes the empirical relationship between mean RT and stimulus intensity as a power function. The FP is the time from the start of the trial until the appearing of the stimulus, which usually has the effect of increasing the mean RT. In this work we tested the Rate Model by using a *simple* reaction time design. Twelve subjects were tested for three blocks of 250 trials each. We varied the FP time and the luminance of the stimuli. The participants were asked to press a button as soon as they saw the stimulus (a luminous circle on a black background). We used 3 FP conditions (0.6, 1, and 2.4 seconds) and 5 luminance levels (0.42, 0.71, 1.21, 2.06, 3.50 cd/m^2). As expected, the relationship between RT and stimulus intensity followed Piéron's Law. We also found that longer FP induced longer mean RT. Fitting a different Piéron's function for each FP condition showed that the effect of the FP was a shift in the Piéron's curve. Overall, the distributions were approximately normal in the rate domain. The goodness of fit increased with the long FP conditions, whereas the different luminance levels did not have any impact on the normality shape of the distributions. We adapted the original model to account for simple RTs, and were able to fit the model to our data to take into account Piéron's Law, the FP effect of shifting the Piéron's curve, and maintain approximately normal distributions in the rate domain (reciprocal RT).Our main claim is that the goal of the human decision process in simple decision tasks is to maximize the rate of reward. We discuss the connections between our model and other recently proposed models for simple RT tasks.

Keywords: *Piéron's Law, Foreperiod, Reward Rate, Simple Reaction Time*

---

1. Introduction

The measurement of reaction/response times (RTs) has been a major paradigm in the study of human behavior and decision-making for over a century (Luce, 1986). The two main experimental designs are *simple* and *choice* RTs tasks. In the latter, a participant is asked to make a choice between two or more stimuli according to experimental instruction after stimuli onset . In the simple RT (SRT) task the participant is asked to respond to the stimulus onset, usually as fast and/or as accurately as possible. Most models have focused on how RT distributions and accuracy depend on experimental parameters, and have assumed a stochastic accumulation of evidence towards one alternative or another (Ratcliff and Rouder, 1998). An alternative approach is examine the strategy that participants use to decide when to stop collecting information and produce a response (Harris et al., 2014). We have found that in a *choice* RT task, the rate of response (1/RT) followed approximately a normal distribution, as has also been previously found for saccades (Carpenter, 1981; Carpenter and Williams, 1995). We therefore proposed that RTs maximized the rate of response. This model also makes some predictions about the effect of the stimulus intensity and the foreperiod time (the time from the start of the trial until the occurrence of the stimulus, FP). In this study we explored the rate distributions in a *simple* RT task to see if they are also near-normal. We also varied stimulus intensity and the foreperiod time to test the predictions of the rate model.

It has been known for long time that the intensity of the stimulus (*I*) affects the mean RT (mRT) according to a power function with negative exponent plus a constant, and has become eponymously known as Piéron's Law (1952): $mRT=\alpha I^{-\beta}+\gamma$, where α is a scaling factor, β is a parameter depending on

the type of stimuli used and γ is the asymptotic value reached when $I \rightarrow \infty$. Piéron's Law seems to hold for *choice* as well as *simple* RT (Pins and Bonnet, 1996; van Maanen et al., 2012). FPs provide a cue to the time of stimulus onset, and are known to affect RTs. When the FP is constant, RTs increase with FP. When the FP is variable, however, the relationship becomes more complicated: the RTs are generally longer, but also depend on the particular FP distribution (Niemi and Näätänen, 1981). We investigated this relationship by using an exponentially distributed FP, which does not give any information about the occurrence of the stimulus at time t given that it did not occur before.

2. Methods

2.1 Participants and Stimuli

12 participants (5 females, 7 males) took part in the experiment. All participants had normal or corrected-to-normal vision and no known neurological conditions. All testing were conducted under constant levels of illumination. The stimuli were presented binocularly on a computer monitor. Participants were positioned 57cm from the monitor. The stimuli were luminous circle (2cm) shown on a dark background (0.31 cd/m^2). The five levels of luminance followed an approximately geometrical series (0.42, 0.71, 1.21, 2.06, 3.50 cd/m^2). Each participant was tested in three blocks. During each block, the 5 possible stimuli were presented randomly. We used three different FP conditions: short, medium and long. Each FP condition consisted of a minimum waiting time (0.4s) added to a random tiem drawn from an exponential distribution. FPs longer than 20s were excluded. The mean of the exponential distribution was 0.2s for the short condition, 0.6s for the medium condition and 2s for the long condition, leading to three different FP conditions with mean of, respectively, 0.6s, 1s and 2.4s. The mean FP time was kept constant for each block. The FP conditions were counterbalanced across participant.

2.2 Procedure

A white cross (2 cm) was presented on a black background at the centre of the screen. This was used as a fixation point. After 1 second, the cross disappeared and a white dot (1mm) appeared at the centre of the screen and remained on the centre during the whole trial. The disappearance of the white cross constituted the warning signal for the start of the FP. After the FP, the stimulus appeared and stayed on the screen until the participant pressed the space bar. The participants were asked to press the spacebar as soon as they detected the stimulus. If the participants pressed the spacebar before the stimulus appeared (anticipation response), a negative auditory feedback was provided and the trial was repeated. Otherwise a positive auditory feedback was provided after the spacebar was pressed. We recorded 50 RTs for each luminance level. Thus, we recorded 250 non-anticipatory RTs for each session, and a total of 750 non-anticipatory RTs for each participant. A training session consisting of 20 trials with a short FP was performed by each participant before the start of the experiment.

3. Results

We calculated the median reaction times (mdRT) across all the 12 participants for each stimulus intensity and FP condition. The median of the reciprocal of a random variable is the same as the reciprocal of the median (the mean does not have this property). As expected, the mdRT decreased with the increase of stimulus intensity ($F_{4,44}$=3.68, P=0.013). The mdRT increased with longer mean FP condition ($F_{2,22}$=10.24, P<0.001), as shown in Figure 1, left. No interaction was found between FP and stimulus intensity. We fitted a different Piéron's Law function for each FP condition using the method of least squares. The estimated parameters are shown in Table 1. Figure 1, right, shows the relationship between the mdRT and FP condition for each luminance level. In this case, the relationship appeared to be logarithmic. This was confirmed by a minimum least square fitting which resulted in high values of R^2 (Table 2).

Figure 1. On the right, the aggregated mdRT for all the 12 participants against the 5 luminance levels of the stimuli, for the 3 FP conditions. On the right, the same mdRT against the 3 FP condition for each one of the 5 luminance levels.

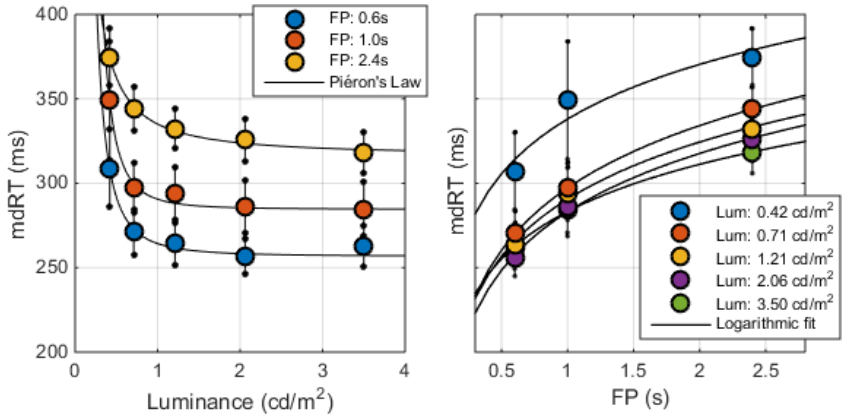


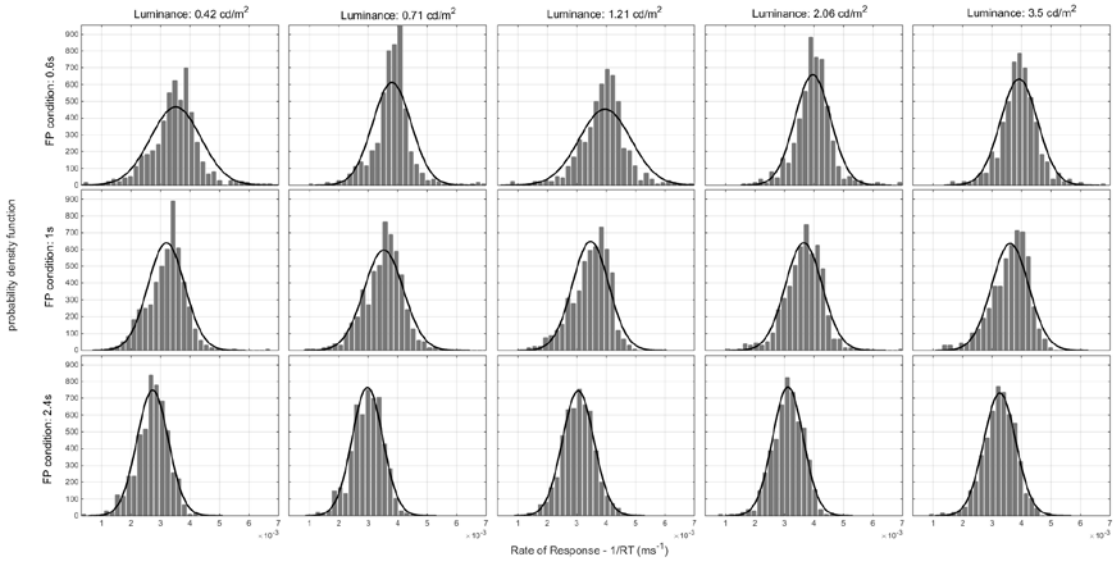
Table 1. Estimated parameters for the Piéron's Law ($mdRT=\alpha I^{\beta}+\gamma$) for each FP condition.

	α	β	γ	R^2
FP=0.6s	9	2.02	254	0.96698
FP=1.0s	6	2.72	281	0.98688
FP=2.4s	20	1.25	317	0.99354

Table 2. Estimate parameters for the relationship between FP and mdRT ($mdRT=k\ log(FP)+c$) for each luminance level.

	0.42 cd/m^2	0.71 cd/m^2	1.21 cd/m^2	2.06 cd/m^2	3.50 cd/m^2
k	46	53	48	49	40
c	338	297	290	283	283
R^2	0.92	~1	0.9938	0.9936	0.9985

Figure 2. Distribution of the rate of response for each condition, aggregated across the 12 subjects using the standardization method.



We analysed the distribution of RTs in the rate domain (1/RT) using the standardization method (Harris et al., 2014). We grouped the distributions for all the participants. The result is shown in Figure 2. We fitted the Normal distribution for each one of the 15 cells in the rate domain (3 FP condition x 5 different luminances) using a maximum likelihood method approach (black solid line)

4. Discussion

4.1 Piéron's Law

The exponent, β, parameter usually found for Piéron's Law in studies in which the luminance of a stimulus is varied is much smaller than the value that we estimated (i.e., Mansfield, 1972, reported β=0.3). We explain this peculiar result by noticing that in our experiment the background luminance was higher than in previous studies (0.31 cd/m^2 against 0.001 cd/m^2), and therefore the contrast stimulus/background in our experiment was much lower than in other studies. High exponents have also been found in Pins and Bonnet (1997, 2000), where the lowest scotopic ranges provided the higher exponent. It is possible that the exponent of the Piéron's function is actually affected by the contrast of the stimuli, and not by the sensory modality as previously assumed. Further studies are required to investigate this point.

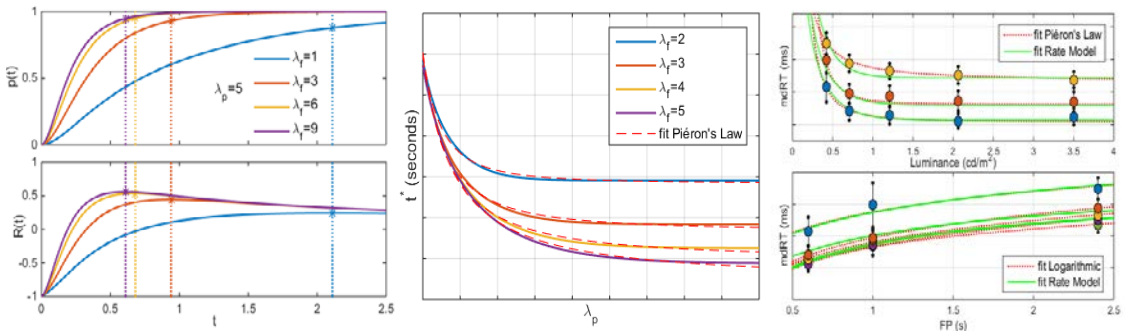
4.2 Distribution shape

We have a previously found in a choice RT task (Harris et al., 2014) that the distributions in the rate domain were approximately Normal. In the present study we use the same approach to verify if the near-Normal distribution shapes can be found in a *simple* RT task with different stimulus intensity and FP conditions. When the mean FP is long the fit with a normal distribution is close to perfect, and the distribution shape does not seem to be affected by the stimulus intensity. However, the fitting gets progressively worse as the mean FP length gets shorter. The FP condition seems to affect the shapes in a non-trivial way that is different for *choice* RTs. In a *choice* RT task the FP does not affect the probability of being correct: even if the participant can estimate when the stimulus is going to occur. In a simple RT task with a variable FP, the flow of time itself can give information about the occurrence of a stimulus, even when an exponential distribution is used. In the next section we will show a modified version of the Reward Rate Model that can take into account the FP effect.

5. The Reward Rate Model for Simple Reaction Time Tasks

The fact that the RT distributions were approximately normally distributed in the rate domain for *choice* RT suggested that the neural computation could be actually carried on in the rate domain (Harris et al., 2014). We proposed a Rate Model in which the rate of reward is maximized. For each trial, we define the gain in subjective utility for a correct response by P>0, and the loss by M>0**.** The expected gain therefore depends on the probability of being correct at time t $\hat{G}(t)=Pp(t)-M(1-p(t))$**.** The expected rate of gain is $\hat{R}(t)=\hat{G}(t)/(t+d)$ where d is the delay across trials. The optimum time to respond (t^*) corresponds to the point when the derivative of $\hat{R}(t)$ is 0, and can be found numerically. The result will depends on the shape of p(t). In our previous work, we used a simple monotonically increasing function asymptotically approaching unity: $p(t)=1-\exp(-\lambda t)$, where λ is a parameters defining how fast p(t) grows (this parameter was used to model different difficulty conditions or different stimulus intensity). However, the results of this study showed us that the FP condition has a peculiar effect on the response. In particular, it changes the γ factor of the Piéron's Law. We show that, by choosing an appropriate formula of p(t), the Rate Model still holds. For the present study we consider a p(t) that depends on both stimulus intensity and FP condition: $p(t)=(1-\exp(-\lambda_p t))(1-\exp(-\lambda_f t))$, which is simple product of two exponent functions, where λ_p and λ_f are two accelerating factors of time, the first depending on the stimulus intensity, the second inversely related with the FP. Figure 3, left and centre panels, show some numerical simulations with several values of λ_p and λ_f. The centre panel shows how the optimum responds is clearly similar to Piéron's Law with different γ depending on the value of λ_f. Assuming that λ_p fluctuates normally across trials, the model can generate distributions that are near-Normal in the rate domain or slightly skewed on the left side, which is similar to what we found in our observations. In Figure 3, right panels, we fit the model to the experimental data by assuming that the mapping between λ_p and λ_f and the physical values of the stimulus intensity and FP condition followed Steven's Law, that is $\lambda_p=\alpha_1 I^{\beta 1}$ and $\lambda_f=\alpha_2 FP^{\beta 2}$. The effect of the stimulus intensity (Figure 3, right-top panel) and of the FP condition (Figure 3, right-bottom panel) can be explained by the model.

Figure 3. Top-left panel: the proposed p(t) with different values of λ_f, which is assumed to be inversely related to FP. Bottom-left panel: the resulting rate of reward (R(t)) function. The vertical dotted line on both figures indicate the optimum time to respond, t that maximized the rate of reward. Centre panel: the resulting shape of the optimum t* with different λ_p parameter, which is assumed to be related to the intensity of the stimulus, and λ_f. Right-panels: experimental data fitted with the Rate Model and Piéron's Law: luminance against mdRT (right-up panel) and FP conditions against mdRT (right-bottom panel).*



6. Conclusion

The effect of stimulus intensity and foreperiod length on RT has been studied for long time. However, the interaction between the two has been largely unexplored. We connected these phenomena with the peculiar fact that RT distributions in the rate domain are normally distributed. This observation holds only partially for simple RT, in which only long FP produced near-Normal distributions in the rate domain. We showed that the Rate Model is also applicable within the framework of simple RT tasks by choosing an appropriate p(t) depending on both FP and RT in order to take into account the FP effect on distribution and on stimulus intensity.

References

Carpenter, R. H., Williams, M. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. Nature, 377(6544) 59-62. doi:10.1038/377059a0

Carpenter, R. H. S. (1981). “Oculomotor procrastination,” in Eye Movements: Cognition and Visual Perception, eds D. F. Fisher, R. A. Monty, and J. W. Senders (Hillsdale: Lawrence Erlbaum), 237–246

Harris, C. M., Waddington, J., Biscione, V., Manzi, S. (2014). Manual choice reaction times in the rate-domain. Frontiers in human neuroscience, 8(June). doi:10.3389/fnhum.2014.00418.

Luce, R. (1986). Response Times: Their Role in Inferring Elementary Mental Organization3. Oxford University Press New York. doi:10.1093/acprof:oso/9780195070019.001.0001

Mansfield, R. J. W. (1973). Latency functions in human vision. Vision Research. 13, 2219-2234.

van Maanen, L., Grasman, R. P., Forstmann, B. U., and Wagenmakers, E.-J. (2012). Pieron's law and optimal behaviour in perceptual decision-making. Frontiers in Neuroscience. 5:143. doi: 10.3389/fnins.2011.00143

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. Psychological Bulletin. 89, 133–162. doi: 10.1037/0033-2909.89.1.133

Piéron, H. (1952). The sensations: Their functions, processes and mechanisms. London: Frederick Muller Ltd.

Pins, D., Bonnet, C. (1996). On the relation between stimulus intensity and processing time: Pieron's law and choice reaction time. Perception & Psychophysics, 58(3) 390-400. doi:10.3758/BF03206815

Pins, D., Bonnet, C. (1997). Reaction times reveal the contribution of the different receptor components in luminance perception. Psychonomic Bulletin & Review, 4(3) 359-366. doi:10.3758/BF03210793

Pins, D., & Bonnet, C. (2000). The Piéron function in the threshold region. Perception & Psychophysics, 62(1), 127-136.

Ratcliff, R., Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. Psychological Science, 9(5) 347-356. doi:10.1111/1467-9280.00067

A NEW PARADIGM FOR INVESTIGATING HUMAN DECISION STRATEGY

Valerio Biscione[1,2], Christopher M. Harris[1,2]
[1]*School of Psychology, Plymouth University, Plymouth, UK*
[2]*Centre for Robotics and Neural Systems and Cognition Institute, Plymouth University, Plymouth, UK*

Abstract

We can divide the decision-making process into two components: 1) the mechanism and neural machinery by which an event is triggered, and 2) the strategy by which the mechanism is employed to solve behavioural/cognitive decision problems. Most of the decision-making models proposed in the past few decades have conflated these two components by assuming that the mechanism constrains and determines behavioural outcomes (RT distributions etc.), as exemplified by the diffusion model. Only more recently has a strategic view been considered (such as maximizing rate or reward). In this work we explore an alternative way to study the strategy component in human decision-making. We introduce a novel paradigm in which a participant is asked to make a decision and gain reward by responding to a task which does not contain any perceptual cue (such as a stimulus onset in a typical RT experiment). We ran several experiments and explored decision strategies proposed in the literature (Bayes Risk, Reward Rate, Reward/Accuracy). We found that participants were appearing to optimize reward rate. We explain some individual differences among participants by referring to a modified reward rate strategy. We also discuss the connection between this task and some classic perceptual tasks. We conclude that focusing on the decision mechanism per se is not enough to understand human decision making, and that a strategic view is necessary.

Keywords: *Optimization, Reward Rate, Speed-accuracy trade-off*

---

1. Introduction

Human decision making is one of the most studied processes in the field of human science, and has drawn attention from many fields: psychology, neuroscience, computer science, mathematics, biology and economics. Most work has focused on answering one or more of the following questions: what are the components of a decision making process? If the decision process is driven by a maximization mechanism (as is usually assumed), what is the objective function that is maximized? Given a certain optimization mechanism and a certain objective function, how good are human in reaching the optimum strategy? In this study, we introduce a new approach to answering these questions.

The most common experimental approach has been the cued choice reaction time paradigm, where a participant is instructed to make a choice among multiple alternatives (usually 2) after the onset of stimulus trial cues. This has led to several models based on three fundamental assumptions: (a) evidence favouring each alternative is integrated over time, (b) the decision is made when enough evidence has accumulated, and (c) the process is stochastic (Bogacz et al., 2006). It is important, however, to distinguish between the decision mechanism and the decision goal. The decision mechanism is the way the perceptual information is gathered and used by the observer (e.g. the rise to threshold mechanism), and the decision goal is the aim of such process (e.g. maximizing reward). Only assumption (b) is related to the decision goal. Indeed, the traditional diffusion model and its variants (Ratcliff and Smith, 2004) have focused on the underlying mechanism, and only recently has the decision goal been considered (Simen et al., 2009; Bogacz et al., 2010). In our view, the distinction between mechanism and goal has been neglected for far too long, mostly because the cued choice reaction time paradigm confounds the two. We propose a new experimental paradigm where the decision goal can be investigated without making any assumption about the underlying mechanism.

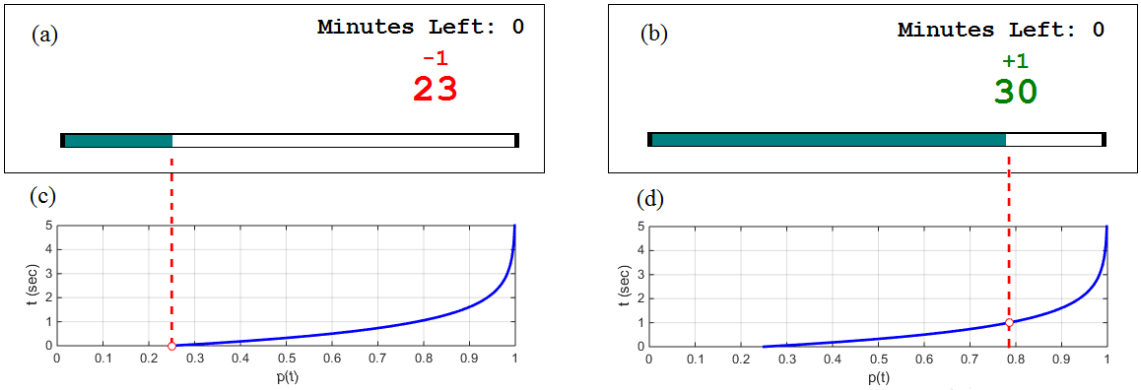
3. The Paradigm

In a cued RT experiment, a speed-accuracy trade-off emerges (Heitz, 2014): the longer a participant waits before responding, the more likely the response is correct. We can define this by a function

p(t), which is the probability that a response made at time t is correct. A speed-accuracy trade-off implies that p(t) is a monotonically increasing function, and has often been construed as a Bayesian accumulation of evidence. However, it is difficult to disentangle how p(t) depends on the perceptual process (anticipatory and post-cue) and how it is used strategically to determine the actual time of response (and hence the probability of being correct). If we want to understand when "enough information is enough" without making any assumption about the process that collects information, we need to hard-code the information sampling mechanism in the experimental design. The rationale is as follows: instead of showing two or more stimuli and having the participant choose the correct one by an endogenous accumulation of information and subsequent increase of p(t), we present p(t) directly to the participant in a well-defined manner. The advantage of this approach is that we have a known p(t), instead of estimating it based on some assumed perceptual model, and it opens up the possibility of manipulating p(t) in a controllable way to understand the decision goal.

To present p(t) directly to the participant, a horizontal gauge is displayed. There are no separate visual cue targets. During a trial the position of the gauge level is moved to the right corresponding to an increasing p(t) predetermined by the experimenter. Zero probability (no reward) corresponds to an empty gauge, and unit probability (reward certain) corresponds to a full gauge. At the beginning of the trial the gauge starts at p(0) (not necessarily zero), and at the end of the trial (t=T), the gauge is at p(T) (figure 1). The participant is fully informed of this relationship, and instructed to press a response button at some chosen time, t, to receive a reward with probability given by the position of the gauge, p(t). We can examine which value of p(t) is chosen by participants.

Figure 1. (a,b) Example of the experimental design used to implement the suggested paradigm, showing how the gauge represents exactly the probability at time t assuming a starting point at 25%. On top there are two instances of the screen shown to the participants. The small red -1 or green +1 is the feedback for the previous trial (point earned or lost). The bigger numbers represent the total score for that condition. The "minutes left" refers to the minutes left for that condition (each condition lasts 3 minutes) (a) the screen at the very beginning of the trial (t=0), (b) after one second (t=1). (c,d)On the bottom there are the p(t) functions and the point corresponding to the probability of being correct at that time, at the beginning of the trial (c) or after 1 second (d).

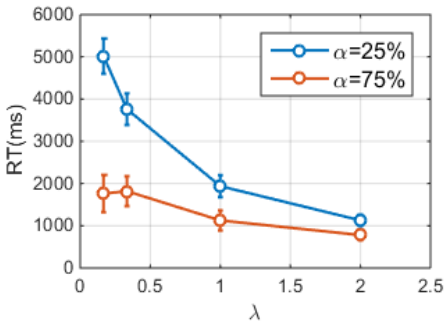


2. Method & Results

To illustrate this new paradigm, we carried out a simple experiment. We defined p(t) to be a monotonically increasing function that asymptotically approached unity according to $p(t)=1-(1-\alpha)\exp(-\lambda t)$, where λ is a parameter that controls ease/difficulty of the trial (increasing λ will make the function steeper, meaning that less time is needed to reach a high probability of correct response), α is the value at t=0 (that is, the probability of getting a correct response at the beginning of the trial). We used 2x4 conditions: two starting points (α=0.25,0.75) and four different speeds of growing of the gauge (λ=0.166, 0.33, 1, 2, higher values indicates faster growing). There was a delay of 0.5s between trials, and each condition lasted 3 minutes regardless of the participants' responses. Twenty participants took part, and were given a training session to familiarise the task. Participants started with 25 points for each condition were instructed to gain as many points as possible.

We averaged the response times across the 20 participants for each of the 2x4=8 conditions and plotted in figure 2. Clearly, for α =0.25, the participants were faster with increasing λ ($F(136,3)=32.48$, $p<0.0001$), as expected. The participants were slower when the starting point is lower (effect of α: $F(136,1)=65.42$, $p<0.0001$)). We fitted the data to 4 different models according to Bogacz et al.'s (2006) classification: Bayes Risk, Reward/Accuracy, Reward Rate, Modified Reward Rate (Harris et al., 2014), and found that the modified Reward Rate provided the best fit.

Figure 2. Mean Reaction Time for all the 20 participants for the two starting point conditions (α) and the four gauge speed conditions (λ).Error bars are ± 1 standard error.



4. Discussion

The participants adjusted their mean response times according to the values of λ and α, and in a way that was consistent with the decision strategy of maximizing reward rate. In varying the parameter λ for α =0.25, the responses were similar to that observed in typical reaction time experiments. For α =0.75, the dependence on λ was much shallower with a local maximum at λ =0.33. To our knowledge, this shape has never been observed in a force choice reaction time experiment, but is consistent with the Modified Reward Rate model (Harris et al., 2014). For the classic choice reaction time paradigm, this suggests that under some conditions (e.g. high α), the expected speed-accuracy trade-off will not be observed. We are now exploring this testable prediction.

In conclusion, we propose a new paradigm in order to investigate human decision making without relying on perceptual accumulation of information models. We have shown that participants were responsive to this new paradigm, and that we can make reasonable and novel inferences about the decision goal. This paradigm can be applied to a great variety of experiments to investigate human decision making, and the results can be easily verified and generalized to more classic studies.

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychological review, 113(4) 700-765. doi:10.1037/0033-295X.113.4.700

Bogacz, R., Hu, P. T., Holmes, P. J., & Cohen, J. D. (2010). Do humans produce the speed-accuracy trade-off that maximizes reward rate? Quarterly journal of experimental psychology (2006), 63(5), 863-891.

Harris C.M., Waddington J., Biscione V., Manzi S. (2014) Manual choice reaction times in the rate-domain. Frontiers in Human Neuroscience 8 doi:10.3389/fnhum.2014.00418

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. Frontiers in Neuroscience. doi:10.3389/fnins.2014.00150

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., & Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. Journal of experimental psychology. Human perception and performance, 35(6), 1865-1897.

Ratcliff, R., Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. Psychological review, 111(2) 333-367. doi:10.1037/0033-295X.111.2.333

A new approach in studying Decision-Making: the EXogenous ACcumulation Task (EXACT)

Valerio Biscione, Christopher M. Harris

Centre for Robotics & Neural Systems and Cognition Institute, Plymouth University, Plymouth, UK

Short abstract.

Human decision-making has been analysed from two viewpoints: 1) the perceptual process that allows observers to accumulate information; and 2) the decision rule that observers employ to decide when to stop collecting information and make a response. In a classic reaction time task the decision rule can only be analysed by assuming a particular perceptual process. We propose a new approach in which the accumulation of information is "hard-coded" in the experiment, so that the researcher can make inferences on the decision rules without having to assume a particular perceptual process. We call this design the "EXogenous ACcumulation Task", because the accumulation of information happens outside the participant's mind. We show experimentally how this paradigm can be used to analyse decision rules. We also compared four different decision rules and show that the modified Reward Rate model fits data better than the Bayesian Risk, Reward Rate, or Reward/Accuracy models. We also show how several features of the response distributions, usually found in classic RT tasks and deemed as side-effects of the perceptual process, are also found with the EXACT and may therefore be explained by the decision rule alone. We hypothesize how Piéron's Law, a psychophysical law that connects speed of response with stimulus intensity, may be the result of the decision rule adopted. We propose several ways in which the EXACT Paradigm could contribute to the understanding of decision-making mechanisms in humans.

Long abstract.

Background. The issue of optimal performance has fascinated researchers for decades. Several experimental paradigms have been used to investigate human decisions and to understand what criterion participants use to decide when to stop accumulating information and make a response. The usual approach consists of 1) collecting data from a classic reaction time (RT) experiment, such as random dot motion discrimination task, 2) assume a particular process

underlying the perceptual mechanism, 3) examine different decision rules. The perceptual mechanism is fitted to the data as well, but rarely more than one process is compared. Most importantly, the predictions of the decision rules depend on the perceptual mechanism itself (such as a diffusion process). We propose a different approach which we called the "EXogenous ACcumulation Task" (EXACT).

The EXogenous ACcumulation Task (EXACT). Every perception process assumes that a participants accuracy increases with time due to an endogenous accumulation of information, which generates a speed-accuracy trade-off. We can define this relationship by a function $ACC(t)$, the probability of making a correct response at time t. To disentangle $ACC(t)$ with the underlying perceptual process, we use an exogenous accumulator perfectly known by the researcher (which is the reason behind the model name and acronym). We display $ACC(t)$ directly to the participant in the form of a visible horizontal gauge. At the beginning of the trial the gauge starts at some point, not necessarily zero, and then the gauge level 'fills up' to the right corresponding to an increasing $ACC(t)$ predetermined by the experimenter. The participant is instructed to press a key to decide when to request a reward with probability equal to aas displayed by the gauge level.

Methods. We chose $ACC(t)$ to be a monotonically increasing function that asymptotically approached unity: $ACC(t) = 1 - (1 - \alpha)exp(-\lambda t)$ where λ is a parameter that controls ease/difficulty of the trial (increasing λ will make the function steeper, such that less time is needed to reach high probabilities of correct response), α is the value at t=0 (that is, the probability of getting a correct response at the beginning of the trial). We used 2x4 conditions: two starting points ($\alpha = 0.25, 0.75$) and four different speeds of the gauge ($\lambda = 0.166, 0.33, 1, 2$). There was a delay of $0.5s$ between trials, and each condition lasted 3 minutes regardless of the participants' responses. Twenty participants took part, and were given a training session to familiarise the task. Participants started with 25 points for each condition and were instructed to gain as many points as possible.

Results. We compared the average response time collapsed across participants for each of the 2x4=8 conditions. For $\alpha = 0.25$ the participants were faster with increasing λ, as expected. The participants were slower when the starting point was lower. We fitted the data to 4 different decision rules (Bogacz et al., (2006): Bayes Risk, Reward Rate, Modified Reward Rate (Harris et al., 2014), and

Reward/Accuracy and found that the Modified Reward Rate provided the best fit. We found a peculiar non-monotonic shape for the $\alpha = 0.75$ condition which is predicted by the Modified Reward Rate but has not been observed in classic RT tasks. We found that the relationship between average responses and λ for the $\alpha = 0.25$ condition followed a power law, commonly found in psychophysics experiments and referred as Piéron's Law.

Conclusion. The EXACT Paradigm shows how it is possible to make inferences about the decision rule without having to rely on a particular model of the perceptual process. In our experiment, participants seemed to be maximizing the rate of reward. The fact that we observed Piéron's Law for the low starting point condition suggests that we are tapping into the same mechanism as in the classic RT experiments. The EXACT Paradigm can be used to better understand decision-making mechanisms and it can be easily compared and generalized to a classic reaction time task.

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychological review, 113(4) 700-765. doi:10.1037/0033-295X.113.4.700

Harris C.M., Waddington J., Biscione V., Manzi S. (2014) Manual choice reaction times in the rate-domain. Frontiers in Human Neuroscience 8 doi:10.3389/fnhum.2014.00418

25th Conference in Subjective Probability, Utility and Decision Making.

Budapest, Hungary. 2015. *(Oral Presentation).*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 10 June 2014
doi: 10.3389/fnhum.2014.00418

Manual choice reaction times in the rate-domain

Christopher M. Harris[1,2]*, Jonathan Waddington[3,4], Valerio Biscione[1,2] and Sean Manzi[2]

[1] *Centre for Robotics and Neural Systems and Cognition Institute, Plymouth University, Plymouth, UK*
[2] *School of Psychology, Plymouth University, Plymouth, UK*
[3] *The WESC Foundation, Exeter, UK*
[4] *School of Psychology, University of Lincoln, Lincoln, UK*

Edited by:
José M. Medina, Universidad de Granada, Spain

Reviewed by:
José M. Medina, Universidad de Granada, Spain
Fuat Balci, Koc University, Turkey

****Correspondence:***
Christopher M. Harris, School of Psychology, Plymouth University, Drake's Circus, Plymouth, Devon PL4 8AA, UK
e-mail: cmharris@plymouth.ac.uk

Over the last 150 years, human manual reaction times (RTs) have been recorded countless times. Yet, our understanding of them remains remarkably poor. RTs are highly variable with positively skewed frequency distributions, often modeled as an inverse Gaussian distribution reflecting a stochastic rise to threshold (diffusion process). However, latency distributions of saccades are very close to the reciprocal Normal, suggesting that "rate" (reciprocal RT) may be the more fundamental variable. We explored whether this phenomenon extends to choice manual RTs. We recorded two-alternative choice RTs from 24 subjects, each with 4 blocks of 200 trials with two task difficulties (easy vs. difficult discrimination) and two instruction sets (urgent vs. accurate). We found that rate distributions were, indeed, very close to Normal, shifting to lower rates with increasing difficulty and accuracy, and for some blocks they appeared to become left-truncated, but still close to Normal. Using autoregressive techniques, we found temporal sequential dependencies for lags of at least 3. We identified a transient and steady-state component in each block. Because rates were Normal, we were able to estimate autoregressive weights using the Box-Jenkins technique, and convert to a moving average model using z-transforms to show explicit dependence on stimulus input. We also found a spatial sequential dependence for the previous 3 lags depending on whether the laterality of previous trials was repeated or alternated. This was partially dissociated from temporal dependency as it only occurred in the easy tasks. We conclude that 2-alternative choice manual RT distributions are close to reciprocal Normal and not the inverse Gaussian. This is not consistent with stochastic rise to threshold models, and we propose a simple optimality model in which reward is maximized to yield to an optimal rate, and hence an optimal time to respond. We discuss how it might be implemented.

Keywords: reaction times, latency, reciprocal Normal, autoregressive integrated moving average (ARIMA), speed-accuracy trade-off, Pieron's law, optimality

INTRODUCTION

Reaction times (response times, latency) (RTs) have been measured and discussed innumerable times since their first measurements in the mid-19th century by von Helmholtz (1850) and Donders (1969). RT experiments are so commonplace that they have become a standard paradigm for measuring behavioral responses, often with scant regard to any underlying process. However, the mechanisms behind RTs are complex and poorly understood. A common view is that RTs reflect processing in the time-domain, where RTs are the sum of independent sequential processes including conduction delays, decision-making processes, and motor responses. We question this very fundamental assumption and consider responses in the rate-domain, where rate is defined as the reciprocal of RT.

One of the most perplexing aspects of RTs is their extreme variability from one trial to the next with some very long RTs, even when the same stimulus is repeated and subjects are instructed to respond as quickly as possible. As exemplified by the saccadic system, why does it take hundreds of milliseconds to decide to make a saccade, when the saccade itself only takes a few tens of milliseconds to execute (Carpenter, 1981)? Moreover, if we accept that point-to-point movements, such as saccades and arm reaching are time-optimal (Harris and Wolpert, 1998), should we not expect the RT also to be optimized? One is then led to wonder how such long response times could be optimal.

DRIFT DIFFUSION MODELS (DDM)

The most popular explanation for the variability of RTs has revolved around the putative mechanism of an accumulator or "rise to threshold" model. A signal, $\rho(t)$, increases (accumulates) in time until it crosses a boundary ("trigger level" or "decision threshold"), $\theta(t)$, whereupon the response is initiated (first-passage time; **Figure 1A**). Typically, $\rho(t)$ is assumed to be a stochastic signal reflecting the accumulation of "information" for or against an alternative until a predetermined level of confidence is reached represented by a constant $\theta(t)$ (Ratcliff, 1978) (**Figure 1B**). A simple reaction time is modeled by a single boundary, and a two-alternative choice task is modeled by

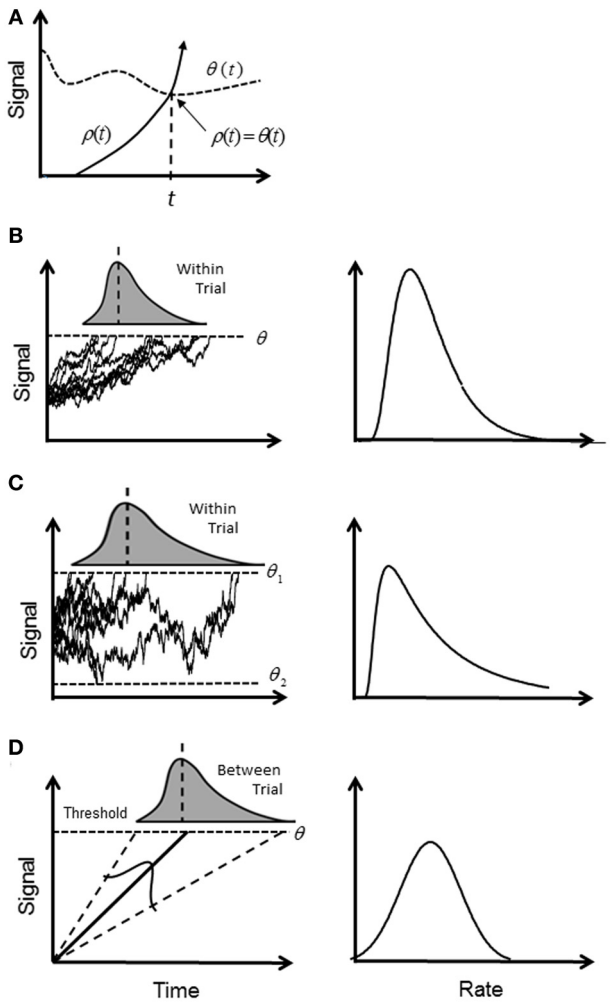


FIGURE 1 | Illustration of accumulator models: (A) general first-passage scheme where a triggered event occurs when a signal $\rho(t)$ first crosses the trigger level $\theta(t)$. Note that crossing is a solution to the equation $\rho(t) = \theta(t)$; **(B)** in the diffusion model $\rho(t)$ increases stochastically and triggers a response when it reaches a constant $\theta(t)$, or "threshold." The signal is assumed to be a Wiener process, and the first passage time is a within-trial random variable (shaded curve) with an inverse Gaussian distribution. In the rate domain (right column) the rate distribution remains positively skewed; **(C)** the diffusion model for two boundaries, where boundary $\theta_1(t)$ determines correct responses, and boundary $\theta_2(t)$ determines error responses. In the rate domain, rate of correct responses remains positively skewed. **(D)** In the deterministic model, $\rho(t)$ increases linearly and deterministically until the threshold is reached. It is assumed that the slope of rise is a between-trial Normal random variable and gives rise to a reciprocal Normal distribution. In the rate domain, rate is distributed with a truncated Normal distribution.

two boundaries. A RT is then first-passage time for one of the alternatives plus any other "non-decision" time such as sensorimotor delays (e.g., Ratcliff and Rouder, 1998; Ratcliff et al., 1999).

Typically, $\rho(t)$ is assumed to drift with a constant mean rate but is instantaneously perturbed by a stationary Normal white noise process (Wiener process), so that within a given trial and with one boundary, the time of crossing the threshold is a random variable with an inverse Gaussian distribution (Schrodinger, 1915; Wald, 1945). With two boundaries, the first passage time for one boundary indicates the decision time for a correct response, and an error response for the other boundary; their probability density functions (pdf's) are computed numerically (Ratcliff, 1978; Ratcliff and Tuerlinckx, 2002) (see **Table 1** for pdf's). For an easy choice task (i.e., high drift rate toward the "correct" boundary), the pdf will approach the inverse Gaussian distribution as error rate become negligible. Although, there are numerous variations on this theme (e.g., Ratcliff and Rouder, 1998, 2000; Smith and Ratcliff, 2004; Bogacz et al., 2006; Ratcliff and Starns, 2013), they share the same basic stochastic rise to threshold decision-making process in the time-domain. It has been recently shown how the pure diffusion process (without variability across trials) has an exact equivalent in terms of Bayesian inference (Bitzer et al., 2014). As shown by Bogacz et al. (2006), the DDM is optimal in the sense that for a given boundary (decision accuracy) the decision is made in minimal time.

Ratcliff (1978) also allowed the mean drift rate to fluctuate between trials with a Normal distribution to reflect "stimulus encoding" variability. This version has often been called the extended DDM, which also includes variability in the starting point of drift, and variability in the non-decision component (Ratcliff and Tuerlinckx, 2002). The extended DDM has been used to describe simple RT experiment (Ratcliff and van Dongen, 2011) and choice RT (Ratcliff, 1978; Hanes and Schall, 1996; Ratcliff and Rouder, 1998; Schall, 2001; Shadlen and Newsome, 2001; Ratcliff et al., 2003, 2004; Smith and Ratcliff, 2004; Wagenmakers et al., 2004; Ratcliff and McKoon, 2008; Roxin and Ledberg, 2008).

Although the multi-parameter extended DDM is claimed to fit observations, a serious problem has emerged from the eye movement literature, when we consider the distribution of the reciprocal of RTs, which we call "rate."

THE RECIPROCAL NORMAL DISTRIBUTION

Investigations into the timing of saccades for supra-threshold stimuli have shown that the frequency distribution of simple RTs (latency) is close to the reciprocal Normal distribution; that is, rate has a near-Normal distribution. Small deviations from true Normal are observed in the tails, but probit plots are typically linear between at least the 5th and 95th centiles (Carpenter, 1981). The reciprocal Normal is not known to be a first-passage distribution for a constant threshold, and is easily distinguished from the inverse Gaussian or the two-boundary pdf. Carpenter has proposed the LATER model in which the rise to threshold is linear and deterministic, but the slope of rise varies from trial to trial with a Normal distribution (Carpenter, 1981; Carpenter and Williams, 1995; Reddi and Carpenter, 2000) (**Figure 1D**). If Carpenter's findings can be generalized beyond saccades, they are equivalent to the extended DDM without fluctuation in the rise of $\rho(t)$ (i.e., no diffusion) and with only one threshold. There is an obvious difficulty in how to explain a deterministic rise to threshold based on a Bayesian update rule, which is inherently stochastic. Moreover, if the rise is deterministic then

Table 1 | Left column: mathematical expressions of the probability density functions (pdf's) for RTs for a single boundary diffusion model, two boundary diffusion model, and the reciprocal Normal.

Time-domain	Rate-domain
Inverse Gaussian $IG(\mu,\sigma^2)=\left(\frac{\mu^3}{2\pi\sigma^2\chi^3}\right)^{\frac{1}{2}}\exp\left[\frac{-\mu(\chi-\mu)^2}{2\sigma^2\chi}\right]$	Reciprocal inverse Gaussian $recIG(\mu,\sigma^2)=\left(\frac{\mu^3}{2\pi\sigma^2\chi}\right)^{\frac{1}{2}}\exp\left[\frac{2\mu^2-\mu/\chi-\chi\mu^3}{2\sigma^2}\right]$
First passage time distribution for the boundary of the two boundaries *a* and *b* for the pure DDM with diffusion constant, *s* $B(\xi,a,b,z)=\frac{\pi s^2}{(a-b)^2}\exp\left[\frac{\xi(a-z)}{s^2}-\frac{\xi^2 t}{2s^2}\right]\sum_{k=1}^{\infty}k\exp\left[-\frac{k^2\pi^2 s^2 t}{2(a-b)^2}\right]\sin\left[\frac{k\pi(a-z)}{a-b}\right]$ see Ratcliff and Smith (2004)	Reciprocal first passage time distribution boundary a of the two boundaries a DDM $recB(\xi,a,b,z)=\frac{\pi s^2}{t^2(a-b)^2}\exp\left[\frac{\xi(a-z)}{s^2}-\frac{\xi^2}{2ts^2}\right]\sum_{k=1}^{\infty}k\exp\left[-\frac{k^2\pi^2 s^2}{2t(a-b)^2}\right]\sin\left[\frac{k\pi(a-z)}{a-b}\right]$
Reciprocal truncated Normal $rectrN(\mu,\sigma)=\frac{\exp\left[-\frac{(1/t-\mu)^2}{2\sigma^2}\right]}{\sqrt{2\pi}\left(1-\phi\left(-\frac{\mu}{\sigma}\right)\right)\sigma t^2}$	Truncated Normal $trN(\mu,\sigma)=\frac{\exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right]}{\sqrt{2\pi}\left(1-\varphi\left(-\frac{\mu}{\sigma}\right)\right)\sigma}$

Φ = *Normal cdf;* ξ = *drift rate;* a = *upper boundary;* b = *lower boundary;* z = *starting point. Right column: equivalent pdf's in the rate (reciprocal RT) domain. See Harris and Waddington (2012) for the mathematical relationship between the two domains.*

the time to reach threshold is known at the outset, and any competition among alternatives can be resolved very quickly—so why wait?

The reciprocal Normal is a bimodal distribution with positive and negative modes. In the time-domain this would imply very large negative RTs, which would require the response to occur long before the stimulus onset and violate causality. Therefore, we need to consider the reciprocal truncated Normal distribution (**rectrN**), (where the Normal rate distribution is left truncated at or near zero; see Harris and Waddington, 2012). The question is what happens at or near zero rate? For easy tasks where RTs are low, the probability of rate reaching zero (i.e., RT approaching infinity) is negligible and the problem might be dismissed as a mathematical nuance. However, for difficult tasks, the probability becomes significant, as we have shown (Harris and Waddington, 2012). A departure from the reciprocal Normal has been reported for saccade latency to very dim targets, but this has been modeled instead as an inverse Gaussian based on a diffusion process (Carpenter et al., 2009). Clarification is needed on what happens when rates are low.

It has long been known that sequential effects occur in manual choice RTs (Hyman, 1953). In sequences of 2-alternative choice RT experiments, RTs may be correlated with the previous trial (first-order) and also earlier trials (high-order). Moreover, this sequential dependency seems to be a function of whether a stimulus is repeated or alternated (Kirby, 1976; Jentzsch and Sommer, 2002). Sequential dependencies cannot be explained by within-trial noise processes, such as the DDM, unless there are between-trial parameter changes (changes in drift rate or threshold values). If we assume a linear dependence on history (autoregressive model) in the rate-domain, then it could in principle lead to convergence onto the Normal distribution via the central limit theorem.

THE RATE-DOMAIN

It is important, therefore, to identify RT distributions, but this is a non-trivial problem. It is difficult to distinguish among highly skewed distributions in the time-domain. The method of moments is infeasible due to poor convergence (the reciprocal Normal has no finite moments; Harris and Waddington, 2012). Maximum likelihood estimation of parameters requires vast amounts of data to distinguish between models (Waddington and Harris, 2012). There is also the problem of under-sampling at extreme values (Harris and Waddington, 2012) which is further exacerbated by the tendency of many investigators to discard "outliers." It is easier in the rate-domain, although large data sets are still needed. Distributions that are less skewed than the reciprocal Normal (such as the inverse Gaussian) remain positively skewed in the rate-domain, whereas the reciprocal Normal does not. Surprisingly, there have only been a few published examples of manual reaction times in the rate-domain (Carpenter, 1999; Harris and Waddington, 2012), and it is conceivable that saccades are somehow "special." For example, express saccades do not appear to have an equivalent in manual tasks. Another important issue is lack of stationarity, where the mean and variance (and higher moments for non-Normal distributions) change over time. Non-stationarity of the mean is particularly troublesome because it smears out the observed distribution making the RT distribution more platykurtic and heavy-tailed. Non-stationarity is more likely in long recording sessions, as subjects become fatigued and bored by the repetitive nature of RT experiments. Using large sample sizes from prolonged recording sessions may be counterproductive.

When a probability density function (pdf) is known in one domain, the pdf in the reciprocal domain can easily be found. However, it is important to recognize this is not true for moments. For example, the mean of the rate distribution is not

the reciprocal of the mean of the RT distribution (Harris and Waddington, 2012). Thus, it is not possible to infer parametric statistics of rate from RT statistics. Raw data are needed. Therefore, our goal in this study was to explore rate-domain analysis in a typical two-choice manual RT experiment. We imposed two tasks (instruction set) and two levels of stimulus difficulty (brightness difference) in order to explore the effects of truncation, and we used autoregression analysis and z-transforms to examine sequential dependency. To minimize problems of non-stationarity, we recorded only modest block sizes (200) from many subjects (24) and collapsed after standardization. We show that rate is indeed near-Normal and not the reciprocal of the inverse Gaussian. Sequential dependency is evident, but not the cause of the near-Normality. In the discussion we propose a rate model as an alternative to first-passage time models.

METHODS

REACTION TIME RECORDING

Subjects were 24 adults aged between 18 and 45 years old selected through the Plymouth University paid participant pool as an opportunity sample. Subjects were naïve to the experimental procedure. Based on self-report, all participants were required to have normal or corrected-to-normal vision with no known neurological conditions. This study received ethical approval from the local ethics committee.

Stimuli consisted of two solid colored rectangles of different luminances arranged horizontally and displayed on a computer monitor (Hanns-G HA191, 1280 × 1024, at 60 Hz). Both rectangles were displayed in the same green color in Red-Green-Blue (RGB) coordinates against a gray background of luminance 37.1 cd/m^2. Each rectangle subtended a visual angle of 5.5 horizontal and 6.6° vertically, and the inner edges were separated horizontally by 9.6°. Viewing distance was 0.5 m. Subjects were instructed to respond to the side with brighter stimulus by pressing the "z" or "2" key. In the easy task (E), rectangle luminances were 37.6 and 131.6 cd/m^2, and in the difficult task (D), they were 37.6 and 37.8 cd/m^2. Calibration was made with a Konica Minolta LS-100 luminance meter. All luminances and ambient room lighting were held constant for all subjects. The luminances in the (E) and (D) tasks were chosen to yield low and high error rates of 1% and 24% for these tasks respectively based on a pilot study. Two task instructions were used and displayed at the beginning of a block. In the "Urgent" (U) task, the instruction was to "respond as fast as possible," and in the "Accurate" (A) task, to "respond as accurately as possible." Each subject was presented with 4 blocks of 200 trials each. Within a block each trial consisted of the same combination of stimulus and task, either AE, AD, UE, or UD. There were 24 different permutations of blocks, and the order was balanced such that each of the 24 subjects had a unique order. We refer to the "easy" tasks as AE and UE, and the "difficult" tasks as AD and UD.

On each trial the subject was prompted to press the space key to commence the trial and a cross appeared in the center of the screen for 500 ms. Subsequently, the two rectangles appeared after a constant foreperiod of 500 ms. For choice reaction time experiments (unlike simple reaction time experiments), constant and variable foreperiods have similar effects (Bertelson and Tisseyre, 1968). We chose constant to avoid introducing additional variability into the decision process (see Discussion). Stimulus onset was also highly salient, even in the difficult tasks, due to the highly visible colored rectangles. The stimuli remained on screen until a response was made or until a time-out of 60 s occurred (see Harris and Waddington, 2012 for a discussion on the importance of a long time-out). For incorrect responses, feedback was provided in the form of a black cross, which remained on screen for 500 ms. A rest break occurred between blocks.

Reaction times (RTs) were measured from the onset of the stimulus presentation and recorded to the nearest millisecond. Rates were computed by taking the reciprocal RT. Taking reciprocals of integer RTs magnifies the effect of the quantization and can lead to artifactual "clumping" and "gaps" in the rate frequency histograms at high values of rate. We eliminated this by using a dithering technique, where we added a uniform floating point random number between −0.5 and +0.5 ms to each RT before taking the reciprocal (see Schuchman, 1964). This has no statistical effect in the time-domain. RTs less than 0.15 s (i.e., rate >6.67 s^{-1}) were considered anticipatory and not analyzed.

MOMENTS

Sample central moments (mean, standard deviation, skewness, and excess kurtosis) and medians were estimated for each block for RT and rate. Note that moments of RT and rate are not reciprocally related, but depend on the underlying parent distribution. However, median rate is the reciprocal of median RT (see Harris and Waddington, 2012).

We also estimated the mean and standard deviation in the rate-domain assuming the underlying distribution was Normal. The underlying mean and standard deviation of the Normal distribution will differ from the sample mean and standard deviation depending on how much of the underlying Normal distribution is truncated. We therefore obtained maximum likelihood estimates (MLEs) of the underlying Normal parameters from each dataset using the *mle.m* function. This function applied a simplex search algorithm to find the parameters that maximized the log likelihood of the probability density function:

$$f(x;\mu,\sigma,a) = \begin{vmatrix} \dfrac{\varphi\left[(x-\mu)/\sigma\right]}{1-\Phi\left[(a-\mu)/\sigma\right]} & a \le x < \infty \\ 0 & x < a \end{vmatrix}$$

where x is the observed rate, μ is the mean of the underlying (untruncated) Normal distribution, σ is the standard deviation of the underlying distribution, $a = 1/60 = 0.0167\ \text{s}^{-1}$, φ is the standard Normal probability density function (pdf), and Φ is the standard Normal cumulative distribution function (cdf).

SEQUENTIAL ANALYSIS

The partial autocorrelation function (PACF) was computed using the *parcorr.m* Matlab function. The first 10 trials on each block were omitted to avoid contamination from initial transients. The coefficients for the first $m = 20$ lags were computed for each block and averaged across blocks. An autoregressive model (AR) was assumed to be of the form:

$$r_n = a_1 r_{n-1} + a_2 r_{n-2} + \cdots + a_m r_{n-m} + u_n \qquad (1.1)$$

where r_i is the response on the ith trial, a_j, $1 < j < m$ are constant weights, and u_i is a stochastic input on the ith trial (negative indices were assumed to have zero weights). The autoregressive weights, a_j and input u_i are unknown and were estimated using the Box-Jenkins maximum likelihood procedure. We used the *estimate.m* function and an autoregressive integrated moving average (ARIMA) model with only an autoregressive polynomial (i.e., no non-seasonal differencing or moving average polynomials). We assumed the distributional form of u_i to be Normal with constant mean and variance.

An AR model can be converted to the equivalent moving average (MA) series using the standard z-transform method. The z-transforms Z(.)of (1.1) is

$$R(z) = a_1 z^{-1} R(z) + a_2 z^{-2} R(z) + \cdots + a_m z^{-m} R(z) + U(z)$$

where $R(z) = Z(r)$, $U(z) = \mathrm{Z}(s)$. This can be viewed as a discrete time MA system with

$R(z) = B(z)U(z)$where the system response of order m is

$$B(z) = \frac{1}{1 - a_1 z^{-1} - a_2 z^{-2} - \cdots - a_m z^{-m}}$$

To find $B(z)$ we took a partial fraction expansion:

$$B(z) = \sum_{i=k}^{m} \frac{\rho_k}{1 - \lambda_k z^{-1}}$$

where λ_i are the roots and ρ_i the residues Taking the inverse z-transform, we then have:

$$r_n = b_0 u_n + b_1 u_{n-1} + b_2 u_{n-2} + \cdots \tag{1.2}$$

where u_k is the stochastic input on trial k and independent of other trial inputs, $b_0 = 1$, and $b_i = \sum_{i=1}^{m} \rho_k \lambda_k^i$, $1 \le i < \infty$, and was computed in Matlab using the *roots.m* and *residue.m* functions. Note that (1.1) and (1.2) describe the same system, but (1.1) is a feedback description, and (1.2) is the feed-forward description. We chose 6 roots, as this encompassed the obviously larger PACF coefficients. The roots were all within the unit circle indicating stability and the existence of a steady-state.

STEADY-STATE TRANSFER

From (1.2) we can relate the pdf of rate (output), $p_r(r)$ to the pdf of the input where u_i are identical independent random variables with pdf $p_u(u)$, $u \ge 0$. From basic probability theory, (Papoullis and Pillai, 2002) the steady-state output pdf is given by the convolution sequence:

$$p_r(r) = \left[\frac{1}{|b_0|} p_u\left(\frac{u}{b_0}\right)\right] \otimes \left[\frac{1}{|b_1|} p_u\left(\frac{u}{b_1}\right)\right] \otimes \left[\frac{1}{|b_2|} p_u\left(\frac{u}{b_2}\right)\right] \otimes \cdots \tag{1.3}$$

where $\otimes$ is the convolution operator. If $p_u(u)$ is Normally distributed then so is $p_r(r)$. If $p_u(u)$ is not Normal then $p_r(r)$ may or may not converge to Normal depending on $p_u(u)$ and the coefficients b_i. We computed (1.3) numerically for the truncated Normal (see Results).

Consider the case where $p_u(0) = c$ where $c > 0$ which corresponds to the case of truncation and when the RT distribution has no finite moments (see Harris and Waddington, 2012). For one term, we have $p_{r,1}(0) = c/|b_0|$. However, with two terms (one convolution) we have $p_{r,2}(r) = \frac{1}{|b_0||b_1|} \int_0^\infty p_u\left(r - \frac{x}{|b_0|}\right) p_u\left(\frac{x}{|b_1|}\right) dx$. For $r = 0$ and $c < \infty$, $p_{r,2}(r) = 0$. Similarly, for all terms we must have $p_r(r) = 0$, so that truncation is lost and the RT distribution will have a finite mean (but not necessarily higher moments).

RESULTS

Subjects' RTs were clearly sensitive to the task and stimulus manipulations, as shown by the example in **Figure 2A** (left column). When stimulus discriminability was easy, RT distributions were brief with low dispersion (AE and UE), but when difficult, they became longer and much more dispersive (AD and UD). In the rate-domain (reciprocal RT) difficulty resulted in a shift toward zero, but the dispersion remained similar (**Figure 2A** right column). For the difficult tasks, the rate distributions appear to approach zero and possibly became truncated. The difficulty was also evident by the number of errors (∼25% in this example).

Similar patterns were seen in all subjects, as can be seen from the plot of medians of RT for all subjects in **Figure 2B.** Again there was much more inter-subject variability for the difficult tasks, but in the rate-domain the variability was more even (**Figure 2C**). Non-parametric testing (Wilcoxon test) showed that the medians differed significantly between the difficult and easy discriminability (AD∪UD vs. AE∪UE: $p < 0.001$), and between task instructions (AD∪AE vs. UD∪UE: $p < 0.001$).

We computed the sample central moments (mean, standard deviation, skewness, excess kurtosis) in the time- and rate-domains (**Figure 3**) for each task for each subject. In the time-domain (left column), the moments were strongly interdependent, as expected from skewed distributions. Standard deviation increased and skewness and excess kurtosis decreased with the mean (note that skewness and kurtosis are normalized with respect to standard deviation). In the rate-domain (right column), however, the interdependence was much weaker (note the difference in ordinate scales).

Because of possible left truncation, we estimated the mean and standard deviation of the putative underlying Normal rate distribution using MLE (see Methods). We set the left truncation to $0.0167\ \mathrm{s}^{-1}$ corresponding to a time-out of 60 s (**Figure 4**). When the sample coefficient of variation (CV) was less than 0.4 (z-score = 2.5; line in **Figure 4**) the MLE estimates (circles) were seen to agree closely with sample moments (crosses). For higher CVs the MLE moments estimates were shifted from the conventional estimates (shown by up-left lines). These shifts in MLE moments are expected from left truncation, and are consistent with, but not definitive of an underlying truncated Normal distribution. Therefore, we next grouped blocks according whether their truncation was severe, "truncated" blocks (CV > 0.4), or negligible, "untruncated" blocks (CV < 0.4).

GROUP DISTRIBUTION

In the untruncated blocks, we standardized the rate for each trial into a z-score based on the ML mean and standard deviation of

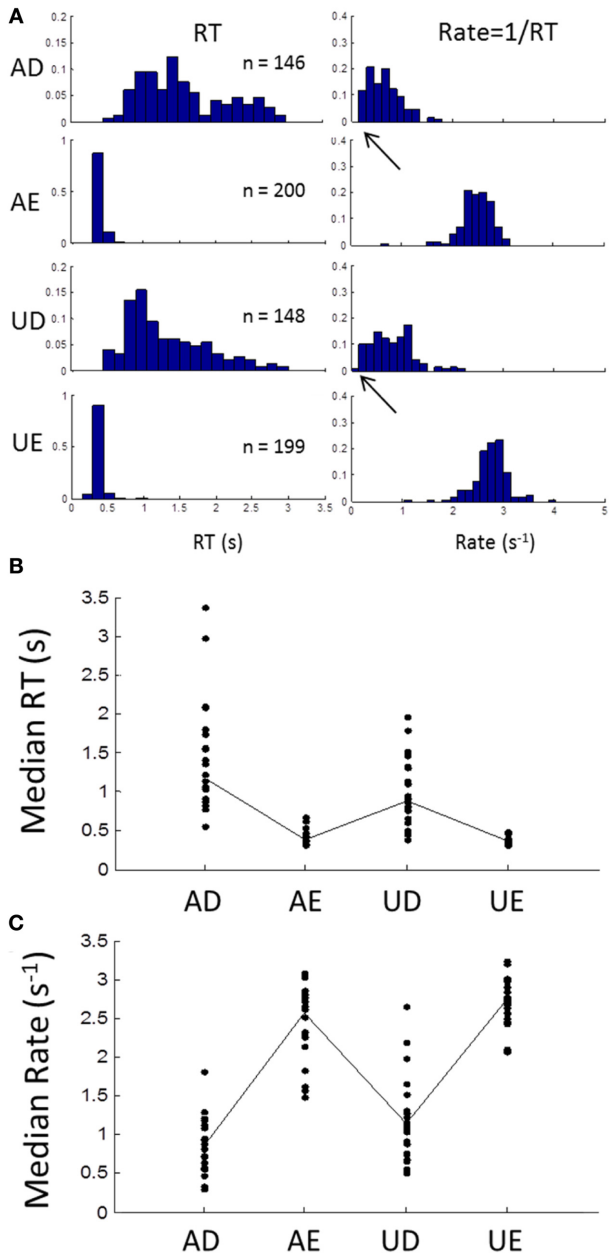


FIGURE 2 | (A) An example of an individual subject's frequency distributions of RT (left column) and rate (right column) for the 4 different blocks (AD, accurate and difficult; AE, accurate and easy; UD, urgent and difficult; UE, urgent and easy; see Methods). In the easy tasks, RTs are brief with few errors (block size was 200 trials). For the difficult tasks RTs are much more variable with about 25% error rate. In the rate-domain, dispersion is similar for all blocks with a shift to lower rates for the difficult tasks. Note that the shift approaches zero (arrows) suggesting possible truncation. **(B)** Median RTs for all subjects showing longer RTs for difficult blocks and more inter-subject variability. **(C)** Same as **(B)** but for median rates showing similar inter-subject variability for all blocks.

its block, and then collapsed all trials into one group. The distribution of the untruncated group was very close to Normal between the 5th and 95th percentile, as seen from the probit plot (**Figure 5A**). There was a slight deviation in the tails. As a check on this method, we created simulated data sets using the true reciprocal Normal distribution with the same ML moments and sample sizes as the empirical data. Carrying out exactly the same analysis, the rate distribution was a perfect Normal—as expected (**Figure 5B**). As a further check, we also simulated the inverse Gaussian. Here there is no truncation issue, so we used sample moments and sample sizes to generate the simulated data. As seen in **Figure 5C**, the reciprocal distribution of the inverse Gaussian is skewed and does not fit the Normal—as expected (Harris and Waddington, 2012). Thus, we are confident that near Normality is not an artifact, but reflects the underlying distribution of the empirical rate distributions.

For the truncated blocks, we standardized as above using the ML mean and standard deviation and collapsed into one group. However, we only considered positive z-scores because any putative truncation would lead to under representation for negative z-scores (we included the one block that had a slightly negative ML mean, see **Figure 3**, but had no discernable effect on the plots when excluded). As shown in **Figure 6A**, the collapsed distribution was close to Normal with a slight deviation above the 95th percentile. Simulation with a true reciprocal Normal showed half a Normal distribution, as expected (**Figure 6B**), and the inverse Gaussian was not close to the truncated Normal (**Figure 6C**). Thus, we conclude that at least the right half of the truncated group are close to Normal, but not the inverse Gaussian. However, this does not address necessarily what happens near zero rate for each block (infra vide).

SEQUENTIAL DEPENDENCY

Temporal effects

The sequence of RTs during a block was clearly not statistically stationary as RTs were typically longer in the first few trials than later. This transient lasted less than 10 trials, after which a steady-state seemed to prevail, best seen by averaging across blocks in the time- or rate-domain (**Figure 7**). The transient was clearly more pronounced for the easy than difficult tasks.

We excluded the first 10 trials of each block in order to examine the steady-state component. The Pearson correlation coefficient between consecutive RTs was 0.20 with 63% of these being significant at $p < 0.05$. In the rate-domain this increased to 0.25 with 76% being significant.

A 1-lag correlation would be expected to lead to autocorrelations with a geometric fall-off at higher lags. Therefore, we examined the partial autocorrelation function (PACF) to explore explicit dependencies up to lags of 20 (see Methods). The PACF of rate was positive and a smoothly decreasing function of lag with no obvious cut-off (**Figure 8A** filled circles). As a check, we shuffled trials randomly within each block and found no significant dependencies (**Figure 8A** open circles). When plotted against reciprocal lag, the PACF coefficients plot was approximately linear (**Figure 8B**; solid circles).

We next considered a stationary autoregressive (AR) relationship of the form: $r_n = a_1 r_{n-1} + a_2 r_{n-2} + \cdots + a_m r_{n-m} + u_n$ (see Equation 1.1 in Methods), where a_i $(1 \leq i \leq m)$ are constant coefficients, u_n is a stochastic input on trial n, which we assumed stationary and Normal, and m is the order of the process (see Methods). We used the Box-Jenkins maximum likelihood

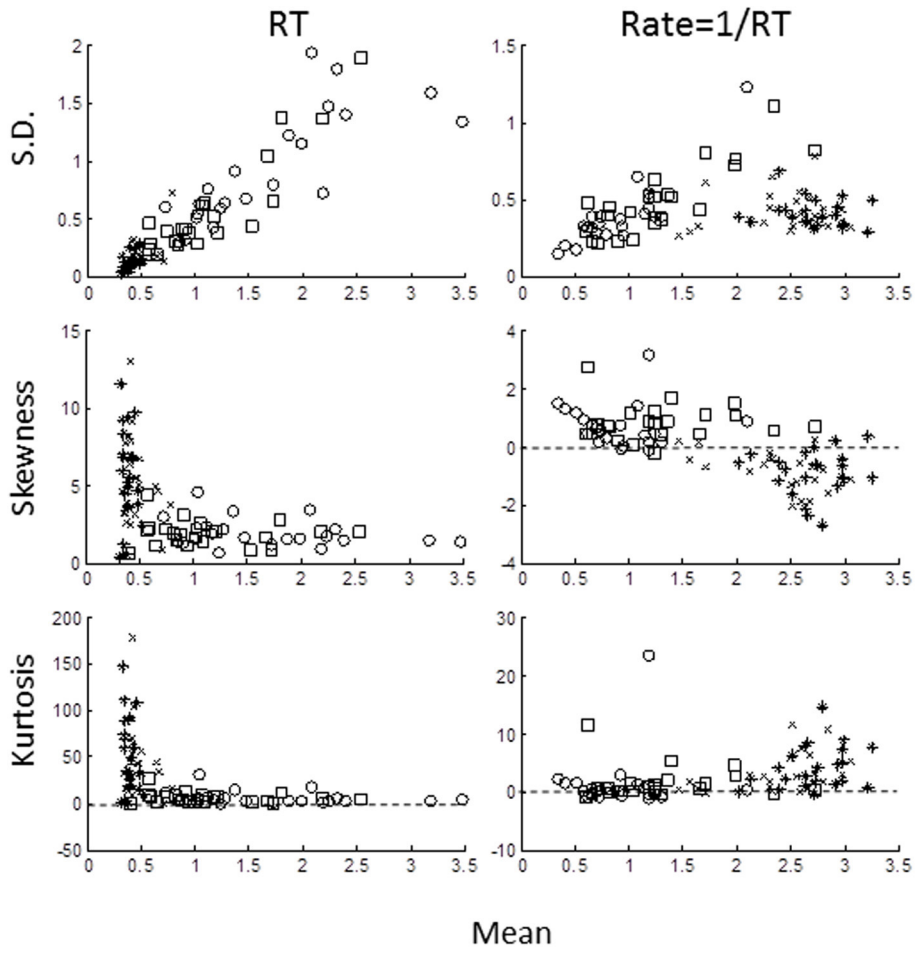


FIGURE 3 | Sample moments for RTs (left column) and rate (right column) plotted against mean for standard deviation (top row), skewness (middle row), and excess kurtosis (bottom row). Each symbol represents the moment for each block for each subject: AD- open circles; AE- crosses; UD- open squares; UE- asterisks). Note different in ordinate scales for RTs and rate moments.

estimation procedure (see Methods) to estimate the a_i for the first 6 lags. We only included "untruncated" blocks (CV < 0.4). Combining all such blocks revealed that only the first 3 weights were significantly different from zero and decreased roughly linearly with reciprocal lag $a_{1,\,2,\,3} = \{0.222,\ 0.104,\ 0.076\}$. The 4th weight $a_4 = 0.016$ was borderline (**Figure 8C**). We also examined the difficult and easy tasks separately, but found negligible difference [AD∪UD: $a_{1,\,2,\,3,\,4} = \{0.212,\ 0.100,\ 0.078,\ 0.016\}$; AE∪UE: $a_{1,\,2,\,3,\,4} = \{0.227,\ 0.105,\ 0.076,\ 0.037\}$]. Henceforth, we used the first 3 weights of the combined tasks.

It is possible to invert the AR process to find the input, since from (1.1) we have $u_n = r_n - a_1 r_{n-1} + a_2 r_{n-2} + \cdots + a_m r_{n-m}$, and the resulting u_n should have no sequential dependency. To test this, we estimated the u_n sequence from each block and re-computed the mean PACF (**Figure 8B** open symbols). Clearly, sequential dependency was eliminated *on average* with a mean lag 1 correlation of 0.032. However, the number of blocks that had a significant lag 1 correlation also dropped from 61 to 10%—which is close to that expected by chance. This implies that most blocks were driven by a similar AR process.

The AR model in (1.1) has a step response which reflects the underlying dynamics behind the steady-state response. It is easily computed (curve in **Figure 8D**) and clearly similar to the empirical average transient response at the beginning of each block (grand average from **Figure 7B**). Thus, the transient response is consistent with the steady-state dynamics.

Using the single-sided z-transform, we converted (1.1) to a moving average (MA) formulation in terms of a discrete series of independent stochastic inputs $u_j\ 1 \le j \le n$ (see Equation 1.2 in Methods): $r_n = b_0 u_n + b_1 u_{n-1} + b_2 u_{n-2} + \cdots$. The weights are the feed-forward impulse response function and are plotted against lag in **Figure 9A**. As can be seen, there is modest but prolonged dependence on input value history implying considerable "memory."

Assuming stationarity, one effect of the sequential dependency is to scale the moments of the input (see Methods). Based on the AR weights, the mean of rate was $\bar{r} = 1.67\bar{u}$. The effect on standard deviation was small $\sigma_r = 1.05(\sigma_u)$, and on higher moments it was negligible. For an untruncated rate distribution, the effect of sequential dependency was to shift the rate distribution to

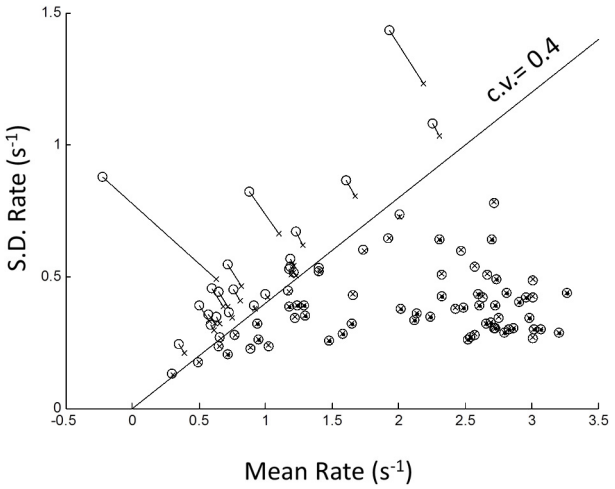


FIGURE 4 | Plot of standard deviation vs. mean of all blocks in the rate-domain. Crosses indicate conventional sample moments (same as top-right panel in **Figure 3**). Circles indicate maximum likelihood estimates (MLE) of same blocks assuming a left truncated Normal. Line is $SD =$ Mean/2.5. To the right of line sample moments coincide with rectrN MLE moments; to the left MLE moments shift to higher standard deviations and lower means (connecting lines).

the right with minimal changes to the shape of the distribution. Thus, we conclude the observed near-Normality of untruncated rate distributions is not a manifestation of the central limit theorem arising from the sequential dependency, but must reflect the near-Normality of the input distribution itself. Therefore, assuming the pdf of the input $p_u(r)$ to be Normal, the output pdf $p_r(r)$ can be computed numerically from the convolution sequence in (Equation 1.3) (see Methods). For an "untruncated" Normal input there is a shift to higher rate with negligible change in variance, as illustrated in **Figure 9B**. For an input truncated at zero, there is not only a shift in the mean, but the sharp truncation at zero is smoothed and eliminated (which can also be demonstrated analytically; see Methods). Remarkably, this smooth shape can also be fit very well by a reciprocal inverse Gaussian (dotted curve) when the tail is excluded (see Discussion).

Spatial effects

Previous studies have shown that mean RT can depend on the sequence of the laterality of previous trials (see Introduction), in particular whether laterality was repeated (**R**) or alternated (**A**). Thus, the sequence **RRRR** indicates that the stimulus and the previous four stimuli were all on the same side (i.e., all left *LLLLL* or all right *RRRRR*), whereas the sequence **AAAA** means that each stimulus alternated sides from the previous (*RLRLR* or *LRLRL*) (note the last symbol is the current trial). Jentzsch and Sommer (2002) examined sequences with 4 lags and showed a significant dependence of RT on a binary weighting of the AR sequence, where **R** was binary "0" and **A** binary "1" (e.g., **RRRR** = 0, **RRAR** = 2, **AARA** = 13, **AAAA** = 16). We used the same scheme for comparison.

For the easy tasks (AE and UE), averaging across all blocks showed a significant dependence on the AR sequence [$F_{(15,\ 645)} = 4.58; p < 0.001$] when all trials in a block were considered. In particular the sequences AARR, RRRA, RRRA, were associated with high RTs (arrow in **Figure 8**), and remarkably similar to Jentzsch and Sommer's results. The inverse pattern was more clearly seen in the rate-domain, with smaller and more even standard errors. For the difficult tasks (AD and UD), there was no significant pattern in the time- or rate-domain.

DISCUSSION

These data clearly show that when the task is easy (AE and UE blocks), RT distributions are close to reciprocal Normal, and not close to the inverse Gaussian distribution. Moreover, we have demonstrated this using practical block sizes ($n = 200$) collapsed across 24 subjects after standardization, unlike previous studies which used very large data sets recorded from only a few subjects. We emphasize that this near-Normality of rate was not an artifact from collapsing across subjects, as this does not invoke the central limit theorem, but simply combines the underlying distributions—as confirmed by Monte-Carlo simulations (**Figure 5B**). We conclude that 2-alternative choice manual RT distributions are very close to the rectrN distribution, similar to the simple reaction experiments with saccades (Carpenter, 1981; Carpenter and Williams, 1995; Reddi and Carpenter, 2000) and the few studies of *simple* manual reaction times (Carpenter, 1999; Harris and Waddington, 2012). In simple RT studies it is necessary to introduce a variable foreperiod to prevent anticipation for the stimulus onset. In choice RT study, a foreperiod may increase "preparedness," but randomization is not essential, as a choice cannot be made with confidence until the discriminative stimulus appears, and Bertelson and Tisseyre (1968) have shown similar effects for constant or random foreperiods in choice experiments. We chose a constant foreperiod to reduce the amount of extrinsic variability introduced into the decision process (see Methods). We can conclude that near-Normality in the rate domain is not a consequence of foreperiod randomization, and by implication presumably neither in simple RT experiments. However, this does not eliminate a possible role of a subject's intrinsic variability in judging foreperiod durations (i.e., Weber's law), and whether or how this affects the rate distribution remains to be explored.

It is difficult to reconcile the rectrN with a pure Wiener diffusion process, where within trial drift noise is Normal (**Figure 1B**), as this would yield an inverse Gaussian distribution in the time-domain, or a reciprocal inverse Gaussian in the rate-domain. Monte Carlo simulation using the reciprocal inverse Gaussian with moments from our subjects did not yield near Normal rates (**Figure 5C**). Ratcliff (1978) considered the compound inverse Gaussian where drift rate fluctuated between trials with another Normal distribution. This would fit the reciprocal Normal if there were no drift noise, which is consistent with Carpenter's LATER model. This strongly suggests that the underlying RT process operates in the rate-domain, rather than in the more intuitive time-domain. It also explains why RTs are so variable—modest symmetric fluctuations in rate can lead to asymmetric and very high changes in RT, especially when rate becomes small as occurs in difficult tasks.

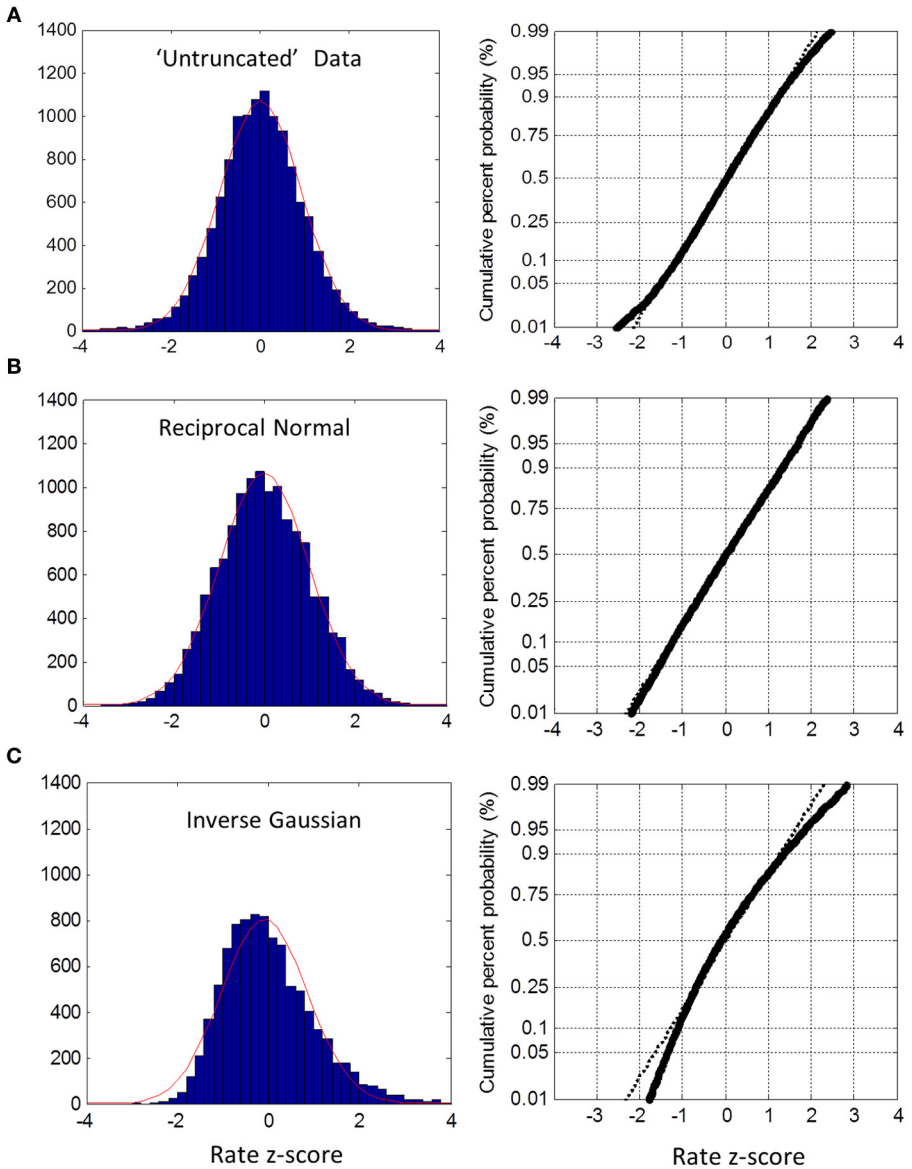


FIGURE 5 | Untruncated group rate histograms (left column) and rate probit plots (right column). (A) Empirical rate from "untruncated" blocks (c.v. < 0.4) showing near Normal distribution over 5–95% interval with slight deviation in the tails **(B)** Simulated data using reciprocal Normal for RT distribution (see text) showing almost perfect Normal rate distribution. **(C)** Simulated data using inverse Gaussian for RT distributions showing obvious deviations from Normal rate.

Temporal sequential dependency among trials has frequently been observed in choice reaction experiments (Laming, 1979). Clearly, any inter-trial correlations affect between-trial fluctuations, but they have been ignored in recent models of RT distributions. Using autoregressive techniques, we have shown explicit dependency of rate output for at least the 3 previous trials, very similar to Laming's original finding in the time-domain. Converting to a MA representation, this "memory" extends even further in terms of stimulus inputs (**Figure 9A**). We also found a transient response at the beginning of each block lasting less than 10 trials, which was similar to the predicted step response of the steady-state dynamics (**Figure 8D**). The simplest explanation is that the rest time between blocks allowed the memory "trace" to decay. However, this needs further exploration since we did not manipulate block intervals, and it was not possible to distinguish between sequential dependencies that are based on absolute time or based on trial number.

Based on moments, the main effect of this temporal dependency was to scale the mean response rate to higher values (i.e., shorten RTs) with little change in variance or higher moments (**Figure 9B**). One could view this as improving signal-noise ratio, or that previous trials/stimuli provide some information about

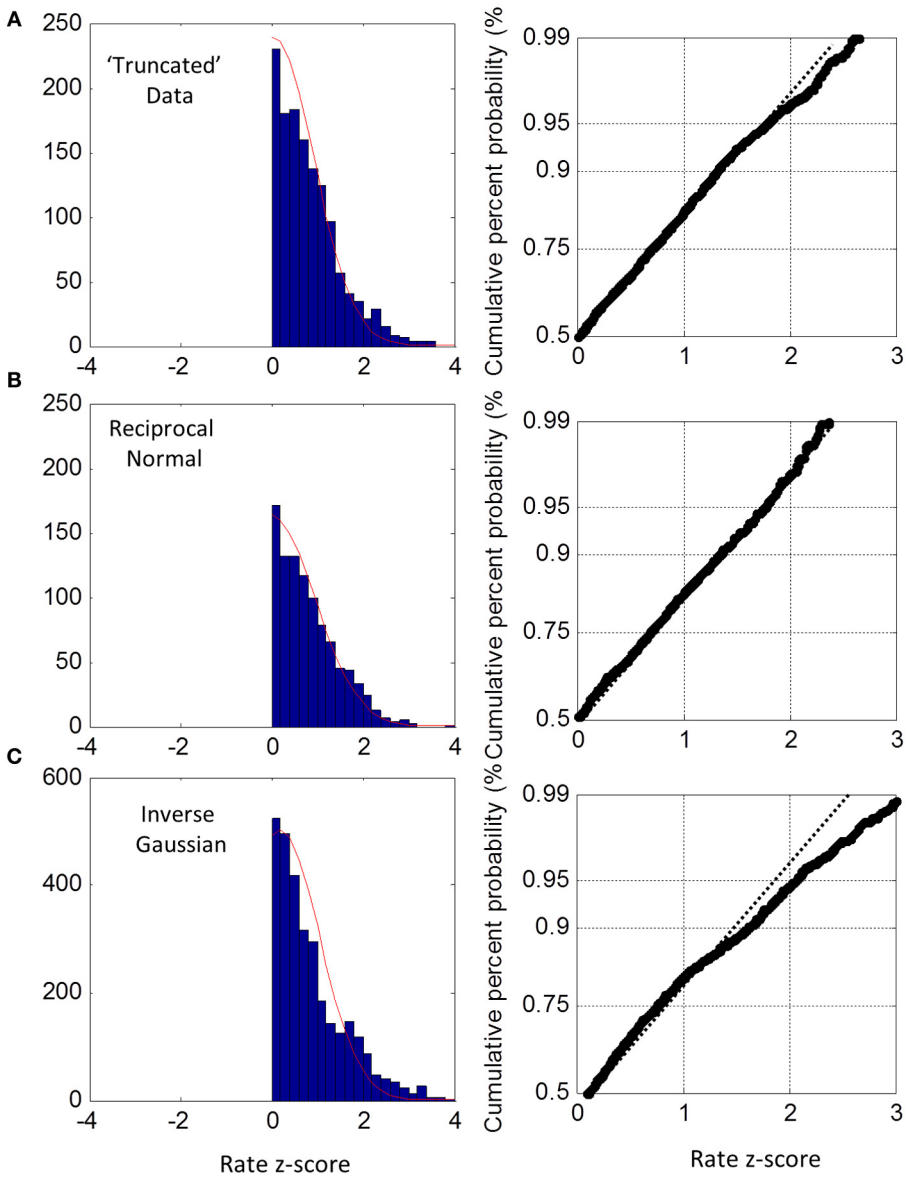


FIGURE 6 | Truncated group rate histograms (left column) and rate probit plots (right column), shown only for positive z-scores (see text). (A) Empirical rate from "truncated" blocks (c.v. > 0.4) showing near Normal distribution over 50–95% percentiles; **(B)** simulated data using reciprocal Normal RTs showing near perfect Normal distributions (as expected). **(C)** Simulated data using inverse Gaussian RTs showing obvious deviations from Normal rate.

the upcoming stimulus (prediction), hence allowing a faster response. Because higher moments are negligibly affected by the MA process, we can also conclude that the temporal sequential dependency does not cause rate to be Normal via the central limit theorem, and we deduce that the input must already be near-Normal.

We also found a sequential dependency that was related to the sequence of stimulus laterality for the easy tasks. Using Jentsch and Sommer's binary weighting system, we found a remarkably similar result to theirs for the easy tasks with **RRRR** and **AAAA** having the highest rates (shortest RTs) and **AAAR**, **RRRA**, **ARRA** having the lowest rates (longest RTs) (**Figure 10**). The weighting scheme of Jentsch and Sommer's extends backward for 4 lags and assumes binary (power function) weighting. From the temporal viewpoint, our results suggest that the 4th lag is questionable and that weightings should follow an approximately hyperbolic decrease. Using this scheme, the dependency becomes even more pronounced (not shown). It is tempting to argue that the temporal and spatial dependencies are manifestations of the same process. Jentsch and Sommer have assumed the dependency reflects a decaying memory trace, as this would explain why higher-order dependencies tend to be weaker when the trials are

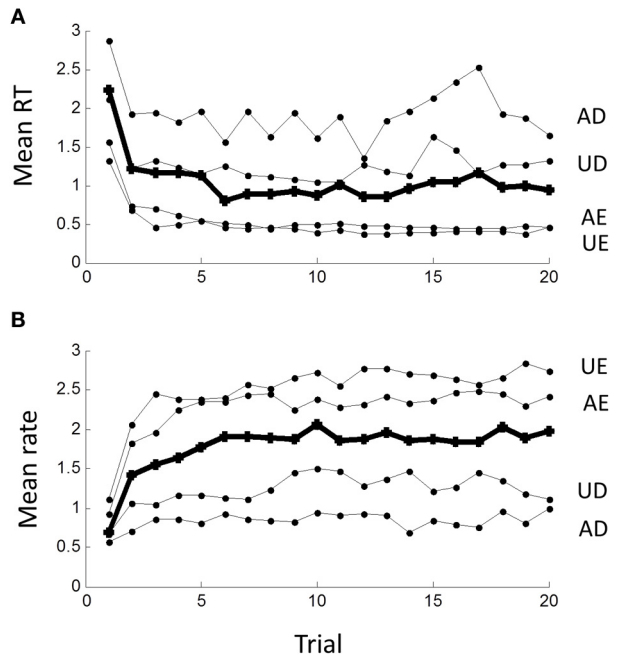


FIGURE 7 | Non-stationarity of responses in (A) the time-domain and (B) the rate-domain. Means are computed across all subjects for the first 20 responses in each block; grand mean across conditions shown by thick line. Note initial transient lasting less than 10 trials, which is more pronounced for the AE and UE blocks.

longer in absolute time. Indeed, we found that the spatial dependency was absent in the difficult tasks (**Figure 10**). Surprisingly, the temporal dependency was still present and virtually identical to the easy task AR process. The reason for this is unclear at present, but suggests that temporal and spatial dependencies can be dissociated.

We emphasize that we have examined sequential dependency in the rate-domain. In the rate domain, a sequence of responses is a well-behaved stochastic process because of its near-Normality, and this permits the wide range of standard analysis techniques (moments, autocorrelations, spectral analyses, etc). In the time-domain this is not necessarily the case because taking the reciprocal of rate is a non-linear operation. Trials with low rates become disproportionately magnified in the time domain, which can lead to "spikes" with very long RTs. In particular, there is the possibility that artefacts may arise in power spectra as these spikes have high spectral energy, and we advocate caution interpreting power spectra based only on time-domain analyses (e.g., 1/f noise: Thornton and Gilden, 2005) subject to further exploration.

TRUNCATION

Strictly, the Normal distribution has infinite extent and includes zero and negative rates, but this is not possible in RT experiments, so we need to consider the left-truncated Normal and the corresponding reciprocal truncated Normal (Harris and Waddington, 2012). We observed that when the task became more difficult (AD and UD), there was a leftward shift of the rate distribution (i.e., longer RTs) (**Figure 2A**) suggesting that left-truncation

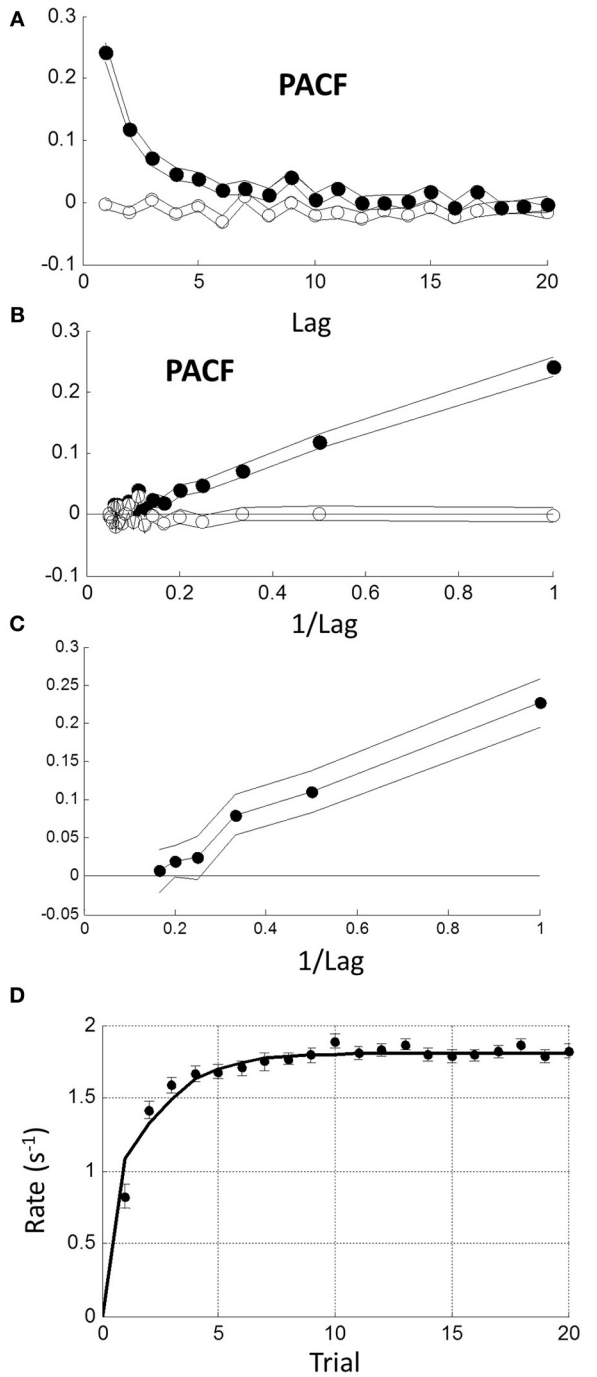


FIGURE 8 | Sequential dependency based on blocks without transients (first 10 trials omitted). (A) Mean partial autocorrelation function (PACF) of all blocks (filled symbols) showing smooth decay. Lines are ± 1 standard error. Open symbols show PACF for the same data after random shuffling leaving no sequential dependency. **(B)** PACF is plotted against reciprocal of lag showing a roughly linear increase (filled symbols). After de-correlation (see text) PACF coefficients become negligible (open symbols). **(C)** Maximum likelihood estimation of autoregressive coefficients (Equation 1.1) using the Box-Jenkins methods (see Methods) showing linear increase with reciprocal lag. **(D)** Comparison of step response function of autorgressive model (solid curve) with observed initial transient from grand mean in **Figure 7B**.

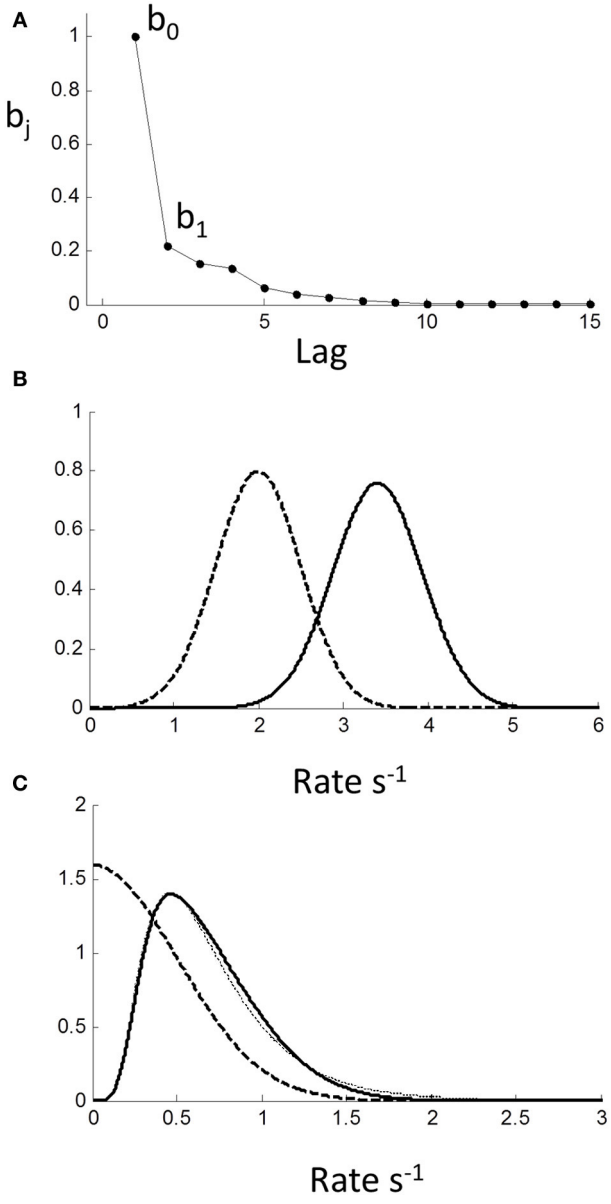


FIGURE 9 | Moving average (MA) model (Equation 1.2) computed from autoregressive coefficients using z-transform method (see Methods). **(A)** MA coefficients show extended dependency on lag indicating memory of input. **(B)** Effect of MA on a Normal input distribution with minimal truncation. Input ($\mu = 2, \sigma = 0.3$) (dashed curve) is shifted to higher rate (solid curve) with little change in shape; **(C)** Input is Normal ($\mu = 0, \sigma = 0.3$) truncated at 0. Note truncation is eliminated by smoothing. Resulting pdf could be mistaken for a reciprocal inverse Gaussian distribution (dotted curve).

may have occurred. Because moments are sensitive to truncation, we used MLE to find the underlying Normal that fitted each block the best, and this showed that truncation was occurring (**Figure 4**). Collapsing across these subjects showed that the untruncated right half of the distribution was also very close to Normal (**Figure 6A**). This is a novel finding, and is evidence that task difficulty can lead to truncated Normal rate distributions. This has not been considered in previous models but has some far-reaching implications.

Truncation leads to very long RTs, which could theoretically approach infinity. Such responses would not usually be observed because either the experimenter imposes a maximum trial duration (time-out), or because the experiment is of finite duration in time or in number of trials. Thus, practically, rate will appear bound at some non-zero minimum, depending of the experimental design (see Harris and Waddington, 2012 for further discussion). For easy tasks, this will have minimal effect since long RTs are rare, but as the task becomes more difficult, the effect of truncation becomes increasingly important.

Interestingly, it has been proposed that the latency distribution of saccades departs from reciprocal normal for low stimulus contrasts, and that the inverse Gaussian is a better model (Carpenter et al., 2009). However, could this instead be due to truncation of the reciprocal Normal? Consider the theoretical example in **Figure 9C**, where we have set the rate standard deviation to $0.3\,s^{-1}$ with left truncation set by a mean of 0. The effect of temporal sequential dependency is to smooth out the truncation, which reduces the probability of very long RTs. The resulting pdf could easily be mistaken for the reciprocal inverse Gaussian (**Figure 9C** dotted curve). Thus, in the time-domain, it is plausible that studies using the inverse Gaussian may have overlooked the reciprocal truncated Normal with sequential dependency as a more parsimonious and unifying explanation.

NON-HOMOGENEITY

In this experiment we have used homogenous and stationary blocks, where the same stimuli were used in each trial of a block, and the laterality was random. However, many RT experiments are not homogenous, and the stimulus value changes on trials within a block. Generally, we expect that rate would no longer be reciprocal Normal. We distinguish between discrete and continuous non-homogeneity.

In the discrete case, a block contains a small number of different but known stimuli that are typically randomized or counterbalanced within the block. Assuming independent trials, the observed rate on each trial would then be a single sample from the Normal distribution associated with that stimulus. The overall rate distribution would then be a mixture of Normal distributions depending on the value and relative frequency of each stimulus. Since the stimulus is known on each trial, responses could be segregated and the rate distributions computed. Clearly, any sequential dependency should be reduced before segregation.

The continuous case is more problematic. It typically occurs when task difficulty and/or stimulus value vary on every trial in an unknown way. The rate on each trial can still be considered as a single sample from a Normal distribution, but the mean of the rate distribution (and possibly the standard deviation) are continuously variable leading overall to a compound Normal distribution, which can take on a wide range of positively or negatively skewed shapes. Whether de-convolving a putative Normal distribution is useful remains to be explored on real data.

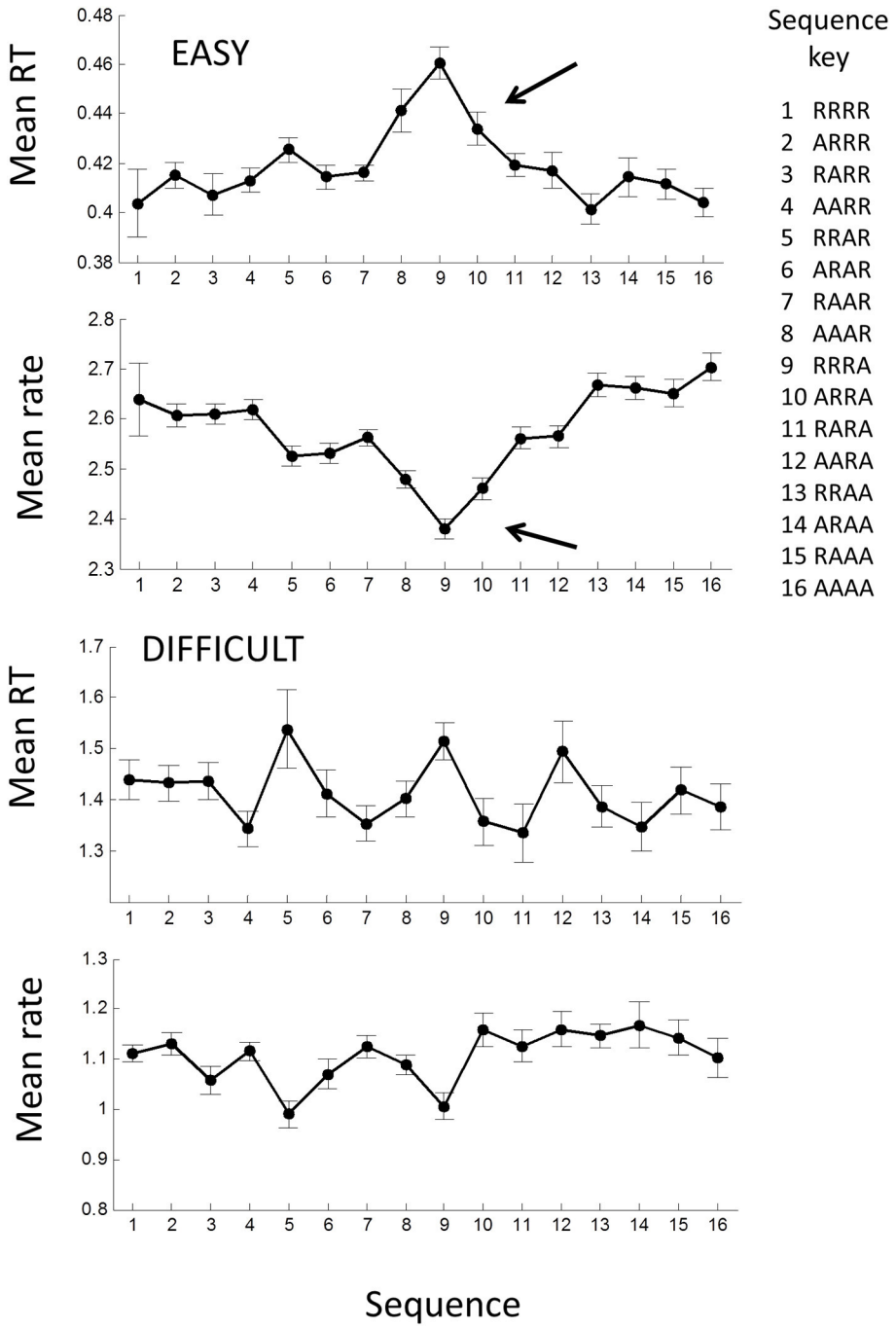


FIGURE 10 | Mean response time and rate plotted against the laterality sequence of previous 4 trials: R, laterality repeated; A, laterality alternated (after Jentzsch and Sommer, 2002). Means were computed across all AE and UE blocks; error bars are ± 1 *within* standard error. Filled symbols show means for all trials in each block; note the significant increase in RT for RRRA sequence (row). A similar picture is seen in the rate-domain.

RATE AND OPTIMALITY

As posed in the introduction, why RTs are so variable and whether, or under what circumstances, they could be optimal are longstanding questions that have been asked or assumed to be answerable by time-domain analysis (e.g., Luce, 1986; Bogacz et al., 2006). However, our and Carpenter's data are highly suggestive that there exists a preferred rate, r^*, for a given set of experimental conditions, and that rate fluctuates according to a Normal random process from trial to trial around r^*. Clearly, modest symmetrical variations in rate can lead to very large and highly asymmetric fluctuations in the time domain, especially when r^* is small—as occurs in difficult discriminative tasks. Also

r^* is easily recognizable as the modal rate, but there is no obvious landmark in the time domain: $t^* = 1/r^*$ does not correspond to the mode in the time-domain. Moreover, the rectrN is a strange distribution without finite moments (Harris and Waddington, 2012), whereas the Normal distribution is a common basic distribution. This strongly suggests that we should be considering rate as the more fundamental variable than RT, even if it seems counter-intuitive.

It seems that if we accept a rise to threshold model, then we require a deterministic drift rate that fluctuates *between* trials with a truncated Normal distribution, as originally proposed by Carpenter (1981). It is conceivable that there is still a stochastic rise to threshold, but it would need to be almost completely masked by the inter-trial variability (this needs future modeling), and rate is still the dominant variable. However, it is important not to conflate proximal with ultimate explanations. At the proximal level, there must be some physiological mechanism for triggering an all-or-none response, and an accumulator process seems physiologically plausible. However, even if true, it only explains how rate could be represented mechanistically, and there is a myriad of ways in which an accumulator could be constructed/evolved as a trigger (e.g., linear vs. curvilinear signal rise, deterministic vs. stochastic signal, fixed vs. variable trigger level; **Figure 1A**). It does not explain why rate is important.

Rate of response may be fundamental for an organism. For example, in the study of natural foraging, it is widely assumed that animals seek to maximize the rate of nutrient intake, rather than quantity *per se*. This has led to the marginal value theorem (Charnov, 1976) which predicts the time spent by animals on patches of food. In the study of animal learning, Skinner introduced his famous cumulative plots as a way of visualizing the stationarity of an animal's rate of response (Skinner, 1938; Ferster and Skinner, 1957). There is an obvious parallel between RT and operant behavior. When a subject presses a button ("operant"), she presumably derives a reward if the button press is a "correct" response, and a loss if "incorrect." The onset of lights acts as a "discriminant" or "conditioned" stimulus that provides information about the probability of reward (Skinner, 1938). It is well known that response times decrease with increasing reward but also increasing intensity of the conditioned stimulus (Mackintosh, 1974). Similarly, numerous studies have shown RTs decrease with increasing reward (Takikawa et al., 2002; Lauwereyns and Wisnewski, 2006; Spreckelmeyer et al., 2009; Milstein and Dorris, 2011; Delmonte et al., 2012; van Hell et al., 2012; Gopin et al., 2013) or increasing stimulus intensity (Cattell, 1886; Piéron, 1914). This leads us to consider the possibility of maximizing expected *rate* of reward or utility as an explanation for our observations (also considered by Gold and Shadlen, 2002).

For each trial, we define the gain in subjective utility for a correct response by $U^+ > 0$, and the loss by $U^- > 0$. Objectively, utility would be maximized by responding to the correct stimulus any time after the stimulus onset. The stimulus value depends on the temporal response of the visual system, and will also increase in time due to any temporal integration and/or Bayesian update of priors. We therefore denote $p(t)$ as the subjective probability of making a correct response given that a response occurs at t (measured relative to some origin; see below). We assume that $p(t)$ is a concave function (**Figure 11A**), where for two alternatives with no prior information, $p(0) = 0.5$.

The expected gain in utility $\hat{G}(t)$ for a response at time t is (curve in **Figure 11B**):

$$\hat{G}(t) = U^+ p(t) - U^- \left(1 - p(t)\right) = \left(U^+ + U^-\right) p(t) - U^- \tag{1.4}$$

It can be seen that expected gain will be negative when $t < t_{\min}$, where $p(t_{\min}) = 1/\left(U^+/U^- + 1\right)$. In this case, it does not pay to respond at all, but there will always be a positive gain as $p(t) \to 1$ and maximized by responding as late as possible. Expected rate of gain is $\hat{R}(t) = \hat{G}(t)/t$. When rate of gain is positive, there may be an optimal time to respond given by $t^* = \underset{t}{\operatorname{argmax}}\, \hat{R}(t)$, which is the solution to:

$$t^* = \frac{\hat{G}(t^*)}{\hat{G}'(t^*)} \tag{1.5}$$

where the dash refers to the derivative with respect to t (**Figure 11C**). The conditions for a positive maximum are complicated, but it occurs under quite broad conditions and is easily visualized geometrically in **Figure 11B**, since from (1.4) the optimum is given by the tangent of $\hat{G}(t)$ that intercepts the origin. Thus, depending on the utility payoff ratio U^+/U^-, and $p(t)$, there is an optimal time to respond. *Responding as quickly as possible is generally suboptimal*—it pays to wait for a specific time to respond.

We can make some general deductions. First, any increase/decrease in the utility payoff ratio, U^+/U^-, will reduce/increase t^* for a concave $p(t)$. Thus, increasing reward will reduce t^*, as empirically observed (*vide supra*). In our experiment, asking subjects to respond accurately as opposed to quickly required "caution" by reducing the ratio and increasing t^* (**Figure 2**).

Faster/slower rise in $p(t)$ will also reduce/increase t^* similar to, but not in precisely the same manner as manipulating payoff. For example, increasing the number of alternatives, n, will reduce $p(t)$ since $p(0) = 1/n$ (given no other prior information) and hence increase t^*. Whether there is a logarithmic relationship between n and $\mathrm{E}[t^*]$ (Hick's law) depends on the precise form of $p(t)$ and remains to be explored. On the other hand, any prior information will decrease the rise-time of $p(t)$ and reduce t^*, as has been reported in some experiments with random foreperiods (see Niemi and Näätänen, 1981).

Stimulus intensity has a strong inverse relationship on t^*, but this depends on $p(t)$. The simplest way to parameterize $p(t)$, is to assume that $p(t)$ depends on a single parameter, ε, that accelerates time so that $p_\varepsilon(t) = p(\varepsilon t)$. We assume that $\hat{\varepsilon}$ is an unbiased estimate of ε and distributed Normally across trials. It follows that $\hat{G}_{\hat{\varepsilon}}(t) = \hat{G}(\hat{\varepsilon} t)$ and $\hat{G}'_{\hat{\varepsilon}}(t) = \hat{\varepsilon}\hat{G}(\hat{\varepsilon} t)$. Then (1.5) becomes

$$\hat{G}'(\hat{\varepsilon} t^*) = \frac{\hat{G}(\hat{\varepsilon} t^*)}{\hat{\varepsilon} t^*} \tag{1.6}$$

so it follows that the optimal solution t^* is given by:

$$t^* = \frac{t_1}{\hat{\varepsilon}} \tag{1.7}$$

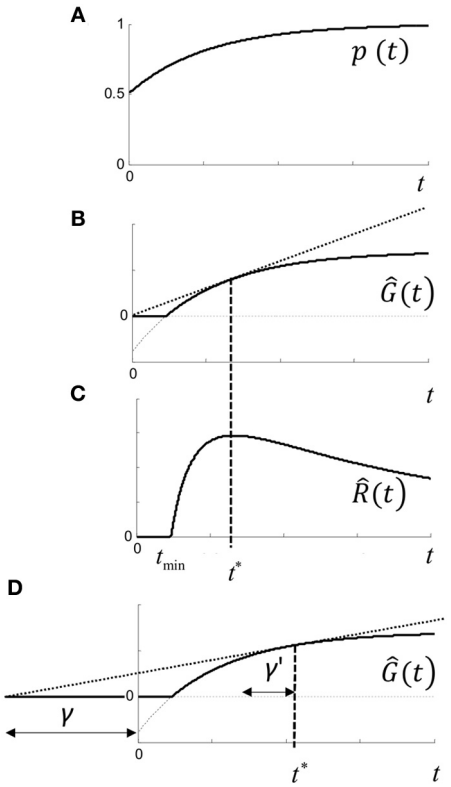


FIGURE 11 | Rate model. (A) $p(t)$ is subjective probability of being correct given a response is made at time t, and is assumed to be concave. Initial value of $p(t)$ assumes guessing with no prior information, and final value is assumes that response will be correct given infinite time. **(B)** $\hat{G}(t)$ is the expected gain in utility (Equation 1.4) for a response made at time t. Note that gain may be negative (i.e., loss) for $t < t_{\min}$ (dashed curve) and no response is optimal. **(C)** $\hat{R}(t)$ is the expected rate of gain in utility (Equation 1.5) which has a maximum at t^*, and can be visualized geometrically as the point where the tangent touches $\hat{G}(t)$ in **(B,D)** shifting the time origin back by γ increases t^* by γ' (see text).

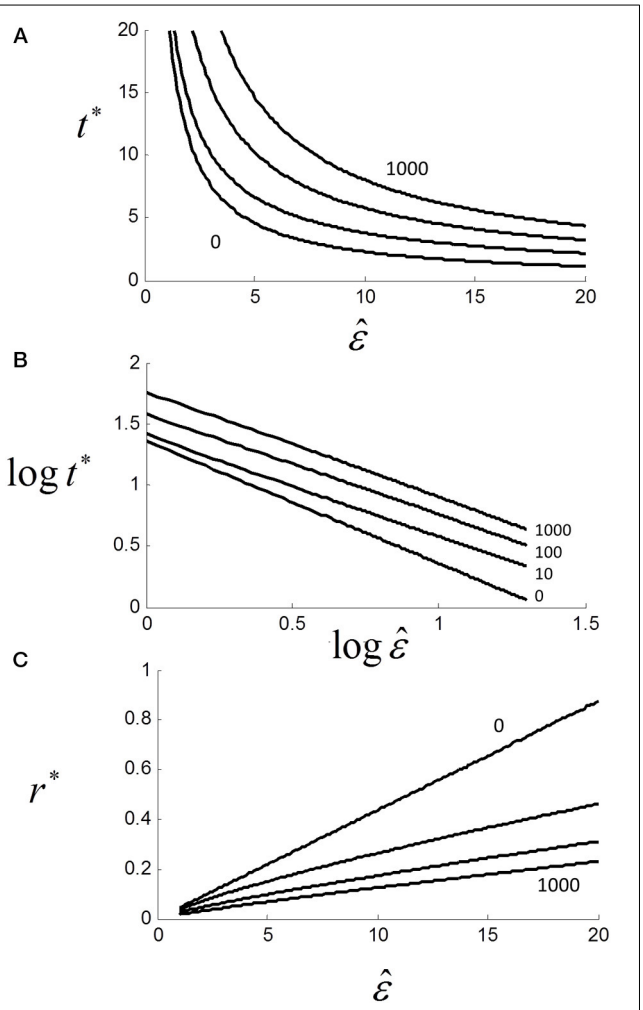


FIGURE 12 | (A) Effect of scaling factor $\hat{\varepsilon}$ on optimal decision time t^* for different non-decision time $T_{ND} = \{0, 10, 100, 1000\}$ (see text). Note that t^* and hence RT increases with T_{ND}, although asymptote is zero (not shown); **(B)** same as **(A)** but on log-log axes (base 10) showing near power function $t^* \approx a\varepsilon^{-k}$ with $k = \{0,\ 0.82,\ 0.83,\ 0.87\}$ and $a = \{25.1,\ 25.1,\ 39.8,\ 63.1\}$ from linear regressions; **(C)** linear plot of optimal rate r^* vs. $\hat{\varepsilon}$. Although strictly a power function, relationship is locally quasi-linear.

where t_1 is the solution to (1.6) evaluated at $\hat{\varepsilon} = 1$. Thus, if each trial is optimized based on the estimate $\hat{\varepsilon}$, then the optimal time to respond is distributed with the reciprocal of the distribution of $\hat{\varepsilon}$ and hence has a reciprocal Normal distribution, as observed.

Since only one reward can occur per trial, we would expect trial duration to be the more relevant epoch for response rate, rather than decision time *per se*. Including an additional non-decision time T_{ND} (foreperiod, sensorimotor delays, etc.) in the computation of estimated rate: $\hat{R}(t) = \hat{G}(t)/(t + T_{ND})$ yields the more general equation for t^*

$$t^* + T_{ND} = \frac{\hat{G}(t^*)}{\hat{G}'(t^*)} \tag{1.8}$$

As shown in **Figure 11D**, including T_{ND} increases optimal response time (relative to stimulus onset). In other words, decision time depends on the amount of non-decision time.

Returning to the parametric model: $\hat{G}_{\hat{\varepsilon}}(t) = \hat{G}(\hat{\varepsilon}t)$, we note that

$$\hat{\varepsilon}\left(t^* + T_{ND}\right) = \frac{\hat{G}(\hat{\varepsilon}t^*)}{\hat{G}'(\hat{\varepsilon}t^*)} \tag{1.9}$$

The solution is not the same as for (1.6), and requires an explicit form for $p(t)$. For the purposes of illustration, we assumed a simple exponential form of $p(t) = 1/2 + \left(1 - \exp\left(-\hat{\varepsilon}t\right)\right)/2$ and plotted t^* against $\hat{\varepsilon}$ with $U^+ = 1$, $U^- = 5$ and parametric in T_{ND} (**Figure 12A**). As can be seen, t^* decays with increasing $\hat{\varepsilon}$ but also increases with T_{ND}. Although we did not manipulate "non-decision" time here, others have shown that increasing foreperiod increases RT in both simple (Niemi and Näätänen, 1981) and choice RT (Green et al., 1983).

For $T_{ND} > 0$, the relationship is still very close to a power law with $t^* \approx a\varepsilon^{-k}$ where $k \approx 0.8$ (**Figure 12B**). In terms of rate, we can see that as T_{ND} increases, r^* decreases but the relationship to $\hat{\varepsilon}$ is still locally close to linear even for very large T_{ND} (**Figure 12C**). Thus, if $\hat{\varepsilon}$ is Normally distributed r^* will also be very near Normal.

If we add sensorimotor delays γ to decision time, then we have $RT = a\hat{\varepsilon}^{-k} + \gamma$ which is clearly similar to Pieron's law: $E[RT] = \alpha I^{-\beta} + \gamma$, where α,β and γ are constants for a given experiment and I is objective stimulus intensity. Piéron's law was originally found for simple RT experiments, but also holds for choice RTs (van Maanen et al., 2012). If we assume that $\hat{\varepsilon}$ is subjective estimated stimulus intensity, then we require $\hat{\varepsilon} \propto I^{\beta/k}$ which is plausible from Steven's power law (Chater and Brown, 1999).

MECHANISM

How optimal rate could be controlled is open to speculation. We can see that the mechanism in **Figure 1A** could act as an equation solver since the time of crossing is the solution of $\rho(t) = \theta(t)$ [more formally: the lowest real positive root of $\rho(t) - \theta(t)$], and when equality is reached, the behavior is triggered in real-time. This can be mapped onto (1.5) in an infinity of ways. A simple possibility is that a deterministic linear rise to threshold behaves as rate-to-time converter (**Figure 1C**). The input $\hat{R}(t)$ is integrated in time to yield a rising deterministic $\rho(t)$ which triggers the response when then a threshold is reached. Gold and Shadlen (2002) proposed that an optimal decision time could be found by an adaptive process (trial-and-error) that varies the threshold. In this case, the distribution of decision times would be given by the distribution of thresholds (for a fixed $\rho(t)$), but this hardly explains why RTs have a near-rectrN distribution. A more parsimonious model would be that the optimal $\rho(t)$ is found for a fixed threshold (i.e., Carpenter's original model). Normally distributed estimates of $\rho(t)$ would then yield RTs with the observed rectrN distribution. It is possible that both threshold and $\rho(t)$ are variable leading to a ratio of distributions for decision time (Waddington and Harris, 2013), although we have no evidence for this in this experiment.

Taking a different perspective, we can draw a correspondence between rate (responses per second) and frequency (cycles per second), and consider control by underlying banks of oscillators in the Fourier domain. It is conceivable that repetitive nature of RT experiments entrain oscillator frequencies, possibly with phase resets from the stimulus onset to allow some degree of prediction. Our observed temporal and spatial sequential dependencies could reflect this entrainment (phase-locking), and the Normal distribution of rate could reflect sampling of subpopulations of oscillators. This is speculative, but not discordant with the known correlation between RTs and alpha brain waves (Drewes and van Rullen, 2011; Diederich et al., 2012; Hamm et al., 2012).

SUMMARY

For 2-alternative manual choice RTs, distributions are close to the reciprocal Normal but not close to the inverse Gaussian distribution. This is not consistent with stochastic rise to threshold models, and implies that between-trial rate (reciprocal RT) is a fundamental variable. There are significant between-trial temporal and spatial sequential dependencies extending back about 3 lags. When tasks become difficult, the rate distributions shift to the left and becomes truncated near zero. We deduced true truncation could not occur due the sequential dependency, but rate distributions are still close to the truncated Normal. Responding to back-to-back sequences of hundreds of almost identical RT trials is not a natural behavior. Nevertheless, it does reflect decision-making when there is time pressure. We propose that when gain in utility is an increasing concave function of time (speed-accuracy trade-off) there emerges an optimal time of response when time is a penalty. We propose that response rate reflects such a process and argue against the longstanding assumption of rise-to-threshold.

REFERENCES

Bertelson, P., and Tisseyre, F. (1968). The time-course of preparation with regular and irregular foreperiods. *Q. J. Exp. Psychol.* 20, 297–300. doi: 10.1080/14640746808400165

Bitzer, S., Park, H., Blankenburg, F., and Kiebel, S. J. (2014). Perceptual decision making: drift-diffusion model is equivalent to a Bayesian model. *Front. Hum. Neurosci.* 8:102. doi: 10.3389/fnhum.2014.00102

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765. doi: 10.1037/0033-295X.113.4.700

Carpenter, R. H., and Williams, M. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*. 377, 59–62. doi: 10.1038/377059a0

Carpenter, R. H. S. (1981). "Oculomotor procrastination," in *Eye Movements: Cognition and Visual Perception*, eds D. F. Fisher, R. A. Monty, and J. W. Senders (Hillsdale: Lawrence Erlbaum), 237–246.

Carpenter, R. H. S. (1999). A neural mechanism that randomises behaviour. *J. Conscious. Stud*. 6, 13–22.

Carpenter, R. H. S., Reddi, B. A. J., and Anderson, A. J. (2009). A simple two-stage model predicts response time distributions. *J. Physiol.* 587, 4051–4062. doi: 10.1113/jphysiol.2009.173955

Cattell, J. M. (1886). The influence of the intensity of the stimulus on the length of the reaction time. *Brain* 8, 512–515. doi: 10.1093/brain/8.4.512

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theor. Popul. Biol*. 9, 129–136. doi: 10.1016/0040-5809(76)90040-X

Chater, N., and Brown, G. D. A. (1999). Scale-invariance as a unifying psychological principle. *Cognition* 69, B17–B24. doi: 10.1016/S0010-0277(98)00066-3

Delmonte, S., Balsters, J. H., McGrath, J., Fitzgerald, J., Brennan, S., Fagan, A. J., et al. (2012). Social and monetary reward processing in autism spectrum disorders. *Mol. Autism* 3:7. doi: 10.1186/2040-2392-3-7

Diederich, A., Schomburg, A., and Colonius, H. (2012). Saccadic reaction times to audiovisual stimuli show effects of oscillatory phase reset. *PLoS ONE*. 7:e44910. doi: 10.1371/journal.pone.0044910

Donders, F. C. (1969). "On the speed of mental processes," in *Attention and Performance II. Acta Psychologica*, ed W. G. Koster (Amsterdam: North-Holland Publishing Company), 412–431. doi: 10.1016/0001-6918(69)90065-1

Drewes, J., and van Rullen, R. (2011). This is the rhythm of your eyes: the phase of ongoing electroencephalogram oscillations modulates saccadic reaction time. *J. Neurosci.* 31, 4698–4708. doi: 10.1523/JNEUROSCI.4795-10.2011

Ferster, C. B., and Skinner, B. F. (1957). *Schedules of Reinforcement.* Englewood Cliffs: Prentice Hall. doi: 10.1037/10627-000

Gold, J. I., and Shadlen, M. N. (2002). Banburismus and the brain, decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308. doi: 10.1016/S0896-6273(02)00971-6

Gopin, C. B., Berwid, O., Marks, D. J., Mlodnicka, A., and Halperin, J. M. (2013). ADHD preschoolers with and without ODD: do they act differently depending on degree of task engagement/reward? *J. Atten. Disord.* 17, 608–619. doi: 10.1177/1087054711432140

Green, D. M., Smith, A. F., and von Gierke, S. M. (1983). Choice reaction time with a random foreperiod. *Percept. Psychophys.* 34, 195–208. doi: 10.3758/BF03202946

Hamm, J. P., Sabatinelli, D., and Clementz, B. A. (2012). Alpha oscillations and the control of voluntary saccadic behavior. *Exp. Brain. Res.* 221, 123–128. doi: 10.1007/s00221-012-3167-8

Hanes, D. P., and Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science* 274, 427–430. doi: 10.1126/science.274.5286.427

Harris, C. M., and Waddington, J. (2012). On the convergence of time interval moments: caveat sciscitator. *J. Neurosci. Methods* 205, 345–356. doi: 10.1016/j.jneumeth.2012.01.017

Harris, C. M., and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature* 394, 780–784. doi: 10.1038/29528

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *J. Exp. Psychol.* 45, 188–196. doi: 10.1037/h0056940

Jentzsch, I., and Sommer, W. (2002). Functional localization and mechanisms of sequential effects in serial reaction time tasks. *Percept. Psychophys.* 64, 1169–1188. doi: 10.3758/BF03194765

Kirby, N. H. (1976). Sequential effects in two-choice reaction time: automatic facilitation or subjective expectancy? *J. Exp. Psychol. Hum. Percept. Perform.* 2, 567–577. doi: 10.1037/0096-1523.2.4.567

Laming, D. (1979). Autocorrelation of choice-reaction times. *Acta Psychologica* 43, 381–412. doi: 10.1016/0001-6918(79)90032-5

Lauwereyns, J., and Wisnewski, R. G. (2006). A reaction-time paradigm to measure reward-oriented bias in rats. *J. Exp. Psychol. Anim. Behav. Process.* 32, 467–473. doi: 10.1037/0097-7403.32.4.467

Luce, R. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York, NY: Oxford University Press.

Mackintosh, N. J. (1974). *The Psychology of Animal Learning*. London: Academic Press.

Milstein, D. M., and Dorris, M. C. (2011). The relationship between saccadic choice and reaction times with manipulations of target value. *Front. Neurosci.* 5:122. doi: 10.3389/fnins.2011.00122

Niemi, P., and Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychol. Bull.* 89, 133–162. doi: 10.1037/0033-2909.89.1.133

Papoullis, A., and Pillai, S. U. (2002). *Probability, Random Variables, and Stochastic Processes. 4th Edn*. NewYork, NY: McGrawHill.

Piéron, H. (1914). Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices. *Année Psychol.* 22, 17–96.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. doi: 10.1037/0033-295X.85.2.59

Ratcliff, R., Cherian, A., and Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *J. Neurophysiol.* 90, 1392–1407. doi: 10.1152/jn.01049.2002

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922. doi: 10.1162/neco.2008.12-06-420

Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356. doi: 10.1111/1467-9280.00067

Ratcliff, R., and Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *J. Exp. Psychol. Hum. Percept. Perform* 26, 127–140. doi: 10.1037/0096-1523.26.1.127

Ratcliff, R., and Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* 111, 333–367. doi: 10.1037/0033-295X.111.2.333

Ratcliff, R., and Starns, J. J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: recognition memory and motion discrimination. *Psychol. Rev.* 120, 697–719. doi: 10.1037/a0033152

Ratcliff, R., Thapar, A., and McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *J. Mem. Lang.* 50, 408–424. doi: 10.1016/j.jml.2003.11.002

Ratcliff, R., and Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. *Psychon. Bull. Rev.* 9, 438–481. doi: 10.3758/BF03196302

Ratcliff, R., and van Dongen, H. P. (2011). Diffusion model for one-choice reaction-time tasks and the cognitive effects of sleep deprivation. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11285–11290. doi: 10.1073/pnas.1100483108

Ratcliff, R., Zandt, T. V., and McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychol. Rev.* 106, 261–300. doi: 10.1037/0033-295X.106.2.261

Reddi, B., and Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nat. Neurosci.* 3, 827–831. doi: 10.1038/77739

Roxin, A., and Ledberg, A. (2008). Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.* 4:3. doi: 10.1371/journal.pcbi.1000046

Schall, J. D. (2001). Neural basis of deciding, choosing and acting. *Nat. Rev. Neurosci.* 2, 33–42. doi: 10.1038/35049054

Schrodinger, E. (1915). Zür Theorie der fall- und steigversuche an teilchenn mit Brownscher bewegung. *Physik. Z.* 16, 289–295.

Schuchman, L. (1964). Dither signals and their effect on quantization noise. *IEEE Trans. Commun. Technol.* 12, 162–165. doi: 10.1109/TCOM.1964.1088973

Shadlen, M. N., and Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936.

Skinner, B. F. (1938). *The Behaviour of Organisms: An Experimental Analysis*. Oxford: Appleton-Century.

Smith, P. L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168. doi: 10.1016/j.tins.2004.01.006

Spreckelmeyer, K. N., Krach, S., Kohls, G., Rademacher, L., Irmak, A., Konrad, K., et al. (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Soc. Cogn. Affect. Neurosci.* 4, 158–165. doi: 10.1093/scan/nsn051

Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., and Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Exp. Brain Res.* 142, 284–291. doi: 10.1007/s00221-001-0928-1

Thornton, T. L., and Gilden, D. L. (2005). Provenance of correlation in psychological data. *Psychon. Bull. Rev.* 12, 409–441. doi: 10.3758/BF03193785

van Hell, H. H., Jager, G., Bossong, M. G., Brouwer, A., Jansma, J. M., Zuurman, L., et al. (2012). Involvement of the endocannabinoid system in reward processing in the human brain. *Psychopharmacology* 219, 981–990. doi: 10.1007/s00213-011-2428-8

van Maanen, L., Grasman, R. P., Forstmann, B. U., and Wagenmakers, E.-J. (2012). Pieron's law and optimal behaviour in perceptual decision-making. *Front. Neurosci.* 5:143. doi: 10.3389/fnins.2011.00143

von Helmholtz, H. L. F. (1850). Über die Methoden, kleinste Zeittheile zu messen, und ihre Anwendung für physiologische Zwecke. *Philos. Mag.* 6, 313–325.

Waddington, J., and Harris, C. M. (2012). Human optokinetic nystagmus: a stochastic analysis. *J. Vis.* 12, 1–17. doi: 10.1167/12.12.5

Waddington, J., and Harris, C. M. (2013). The distribution of quick phase interval durations in human optokinetic nystagmus. *Exp. Brain Res.* 224, 179–187. doi: 10.1007/s00221-012-3297-z

Wagenmakers, E.-J., Steyvers, M., Raaijmakers, J. G. W., Shiffrin, R. M., van Rijn, H., and Zeelenberg, R. (2004). A model for evidence accumulation in the lexical decision task. *Cogn. Psychol.* 48:3. doi: 10.1016/j.cogpsych.2003.08.001

Wald, A. (1945). Sequential tests of statistical hypotheses. *Ann. Math. Stat.* 16, 117–186. doi: 10.1214/aoms/1177731118

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 March 2014; accepted: 21 May 2014; published online: 10 June 2014.
*Citation: Harris CM, Waddington J, Biscione V and Manzi S (2014) Manual choice reaction times in the rate-domain. Front. Hum. Neurosci. **8**:418. doi: 10.3389/fnhum.2014.00418*
This article was submitted to the journal Frontiers in Human Neuroscience.

frontiers in Behavioral Neuroscience

ORIGINAL RESEARCH
published: 03 November 2015
doi: 10.3389/fnbeh.2015.00288

Investigating decision rules with a new experimental design: the EXACT paradigm

Valerio Biscione[1,2]* *and Christopher M. Harris*[1,2]

[1] School of Psychology, Plymouth University, Plymouth, UK, [2] Centre for Robotics and Neural Systems, Plymouth University, Plymouth, UK

In the decision-making field, it is important to distinguish between the perceptual process (how information is collected) and the decision rule (the strategy governing decision-making). We propose a new paradigm, called EXogenous ACcumulation Task (EXACT) to disentangle these two components. The paradigm consists of showing a horizontal gauge that represents the probability of receiving a reward at time t and increases with time. The participant is asked to press a button when they want to request a reward. Thus, the perceptual mechanism is hard-coded and does not need to be inferred from the data. Based on this paradigm, we compared four decision rules (Bayes Risk, Reward Rate, Reward/Accuracy, and Modified Reward Rate) and found that participants appeared to behave according to the Modified Reward Rate. We propose a new way of analysing the data by using the accuracy of responses, which can only be inferred in classic RT tasks. Our analysis suggests that several experimental findings such as RT distribution and its relationship with experimental conditions, usually deemed to be the result of a rise-to-threshold process, may be simply explained by the effect of the decision rule employed.

Keywords: optimal performance, reward rate, speed-accuracy trade-off, perceptual choice, decision rules

OPEN ACCESS

Edited by:
Oliver T. Wolf,
Ruhr University Bochum, Germany

Reviewed by:
Tobias Kalenscher,
Heinrich-Heine University Duesseldorf, Germany
Fuat Balci,
Koc University, Turkey

***Correspondence:**
Valerio Biscione
valerio.biscione@plymouth.ac.uk

Received: *17 June 2015*
Accepted: *12 October 2015*
Published: *03 November 2015*

Citation:
Biscione V and Harris CM (2015) Investigating decision rules with a new experimental design: the EXACT paradigm. Front. Behav. Neurosci. 9:288. doi: 10.3389/fnbeh.2015.00288

INTRODUCTION

Decision-making can be broken down, at least conceptually, into two components: a perceptual process and a decision rule. We define the perceptual process as the mechanism that accumulates information in order to decide between alternative responses. It does not include all the perceptual information an observer is experiencing during a decision/experimental task, but only that which affects the ultimate decision. We define the decision rule as the quantity being optimized during the task, which will therefore establish when enough information has been gathered and a decision can be made. A great amount of work in the decision-making field has focused on the perceptual process implemented in the brain, resulting in several models that can account for a wide variety of data (e.g., LaBerge, 1962; Laming, 1968; Ratcliff, 1978; Busemeyer and Townsend, 1993; Usher and McClelland, 2001; Ratcliff and Smith, 2004; Ratcliff et al., 2004a,b). In most cases, these models assume a noisy accumulation of information toward one of two alternatives, until a decision threshold is reached, whereupon a response is made (e.g., drift diffusion model: Stone, 1960; Laming, 1968; Ratcliff, 1978; Ratcliff and Rouder, 2000; linear ballistic accumulation: Brown and Heathcote, 2008; Ornstein-Uhlenbeck model: Busemeyer and Townsend, 1993). There is, however, a growing interest in the decision rule

per se, which affects how participants set their decision thresholds, that is when to stop collecting information and make a decision (Gold and Shadlen, 2002; Bogacz et al., 2006; Holmes and Cohen, 2014; Moran, 2014).

The usual approach to investigating decision rules consists of (1) collecting data from a classic reaction time (RT) experiment; (2) assume a particular perceptual process, usually the drift diffusion model; (3) based on the perceptual process assumed, test different decision rules (for example, see Simen and Cohen, 2009; Bogacz et al., 2010). This approach has serious drawbacks: the predictions of the decision rule depend on the perceptual mechanism assumed, and there are a variety of possible perceptual mechanisms (e.g., Ratcliff et al., 1999; Ratcliff and Smith, 2004; Smith and Ratcliff, 2004). When considering more than two alternatives, there are even more approaches which differ in behavioral and neurobiological assumptions (Krajbich and Rangel, 2011). This leads to a problem when fitting decision rules: which perceptual process should one assume? Moreover, should a decision rule not fit the data, would that be due to the decision rule itself or to the perceptual process assumed? Hypothetically, this issue could be solved by comparing different perceptual processes with different decision rules, but this is rarely done in practice. Even in this case, it is not possible to know if the sample of perceptual processes tested comprises the one used by human participants. This is a serious limitation when the focus of the investigation is on the decision rule, and not the perceptual process. There is also a second limitation of applying perceptual process in testing decision rules: the commonly used perceptual models are employed in case of fast decisions (average RT < 1 s) (Voss et al., 2013) but their applicability for longer decision remain untested, even though real-life decisions may take several seconds.

We believe that there is a need for a new paradigm which will allow researchers in the decision-making field to reduce the interdependence of the perceptual and decision processes. We propose a new paradigm, called EXACT (EXogenous ACcumulation Task) which allows the decision rule to be investigated without the need of making any assumption about the underlying perceptual process.

The perceptual process affects the decision rule by defining a certain relationship between time of response and accuracy of response: $ACC(t)$, where t is the time taken to make a response. $ACC(t)$ defines the speed-accuracy trade-off: it is the probability of being correct at time t, and it is assumed to increase with time so that the slower the response, the more accurate it is (Heitz, 2014). The particular shape of this function depends on the perceptual model assumed. To understand the decision rule without making any assumption about the perceptual process, we need to hard-code $ACC(t)$ itself into the experimental design. Instead of showing two or more stimuli and having the participant choose the correct one by an endogenous accumulation of information and subsequent increase of $ACC(t)$, the $ACC(t)$ can be presented directly to the participants who are asked to make the decision based on this exogenous $ACC(t)$. We call this paradigm EXACT, because the speed-accuracy trade-off function is presented exogenously to the participant, instead of being an endogenous process. In this task, we are assuming that the perceptual accumulation of information is separable from the decision rule, such that if we replace the perceptual accumulation exogenously, we can observe the decision rule in a meaningful way.

To present $ACC(t)$ directly to the participant, a horizontal gauge is displayed. There are no separate visual targets. During a trial, the position of the gauge level is moved to the right corresponding to an increasing $ACC(t)$ predetermined by the experimenter. Zero probability corresponds to an empty gauge, and unit probability corresponds to a full gauge. At the beginning of the trial the gauge starts at $ACC(0)$ (not necessarily zero). The participant wins X reward tokens with probability $ACC(t)$ when they press a button at time t, and loses Y tokens with probability $1 - ACC(t)$. After the response, a new trial starts. The experimenter can specify a delay between trials. **Figure 1** (top) shows an example of the monitor screen shown to the participants. On the bottom panels the corresponding $ACC(t)$ (an exponential function, as discussed below) is shown (not seen by participants). The axes are inverted so that the horizontal axis corresponds to the direction of the gauge movement.

For an illustrative experiment with the EXACT paradigm, we chose an exponential $ACC(t)$:

$$ACC\,(t) = 1 + (\alpha - 1)\exp\,(-\lambda t) \tag{1}$$

so that:

$$t(ACC) = -\ln\left(\frac{(ACC-1)}{\alpha-1}\right)/\lambda \tag{2}$$

where λ is a parameter controlling how fast the function grows, and α is the value at $t = 0$. In the EXACT paradigm, λ corresponds to the speed of the gauge, and α is the gauge level at the beginning of the trial, that is $ACC(0)$. This choice of $ACC\,(t)$ was based on ease of mathematical tractability, capability of capturing different aspect of the task, and similarity to the $ACC(t)$ produced by widely used models[1]. However, other functions could be used to test decision rules. The main point is that the function is completely known by the researcher and not derived from an assumed perceptual process.

The EXACT paradigm can be easily compared with a classic RT task. The λ parameter controls the speed of the gauge, which can be compared to manipulating the trial difficulty (for example, by changing stimulus intensity, stimulus contrast, etc…). When the task is difficult, the rate of accumulation of evidence is smaller, and $ACC(t)$ grows slower. Similarly, changing α, the gauge level at the beginning of the trial corresponds to the number of alternatives: $\alpha = 0.5$ represents a 2AFC, $\alpha = 0.25$ a 4AFC. With the EXACT paradigm it is also possible to explore more complicated designs, such as $\alpha = 0.75$, which

[1] For example, consider the $ACC\,(t)$ resulting from a drift diffusion model with two boundaries (that is, two alternative forced choice task): $ACC(t) = 1 - \boldsymbol{\phi}\left(-\frac{A}{c}\sqrt{t}\right)$, where A is the drift rate, c is the standard deviation of the process and $\boldsymbol{\phi}$ is the normal standard cumulative function (Bogacz et al., 2006). Notice that this function is monotonically increasing, concave, and asymptotically approaches unity, similarly to our function. However, the function does not allow us to model different starting points, because the drift diffusion model can be only used in case of two alternative forced choice tasks, and the difficulty of the trial is controlled by two parameters instead of one of our function.

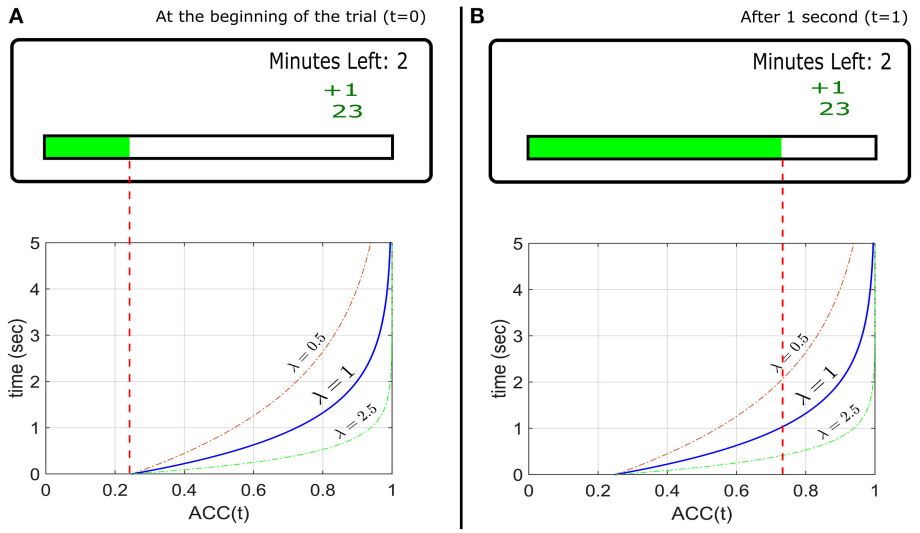


FIGURE 1 | Top Panels: Example of the experimental design used to implement the EXACT paradigm, showing how the gauge represents the probability at time t assuming a starting point at 25% during a hypothetical trial. The small green +1 displays the feedback from the previous trial (point won/lost). The number below represents the total current score for that condition. The "minutes left" refers to the minutes left for that condition. **(A)** Top: the screen as seen by the participant at the beginning of the trial ($t = 0$) and **(B)** after 1 s, assuming $\lambda = 1$. Bottom: the ACC(t) functions with different λs, representing the probability of getting a reward with time. The gauge on the top panels is only referred to the function with $\lambda = 1$. The other functions show how the accuracy would change in time for different λ. For example, with $\lambda = 2.5$, after 1 s, the accuracy is equal to ~0.95 (the gauge would be almost completely full).

corresponds to a task with three correct alternatives out of four possible choices. Conceptual differences from the classic RT task and limitation of the EXACT paradigm are addressed in the discussion (Limitations of the EXACT paradigm).

By knowing the shape of $ACC(t)$, it is straightforward to test possible decision rules. Four decision rules have been proposed frequently in the literature (Bogacz et al., 2006; Simen and Cohen, 2009; Harris et al., 2014; see Materials and Methods for further details):

Bayes Risk (BR). This rule was first used by Wald and Wolfowitz (1948) in proving the optimality for the Sequential Probability Ratio Test. It assumes that decision makers seek to minimize a cost function that is the weighted sum of the time and the error rate, without taking into account the total length of the trial. The optimum response time depend on the ratio between the subjective weights of time and error rate.

Reward Rate (RR). It is defined as the proportion of correct trials divided by the average duration between trials (Bogacz et al., 2006). According to RR, responses should not depend on the payoff matrix used in the task.

Reward/Accuracy (RA). This has been proposed by Bogacz et al. (2006) by formalizing the COBRA theory of Maddox and Bohil, 1998). It is a weighted difference of RR and accuracy, in which the optimum response depends on the punishment/reward ratio and on the total length of the trial.

Modified Reward Rate (*$\mathbf{RR_m}$*). This rule has been proposed by Bogacz et al. (2006) and Harris et al. (2014). It describes a situation where a correct response corresponds to a subjective reward, and an incorrect response to a subjective punishment, and participants are trying to maximize the rate of gain (Harris et al., 2014). The optimum response does not depend on the absolute values of punishment and reward, but only on their ratio. RR only takes into account the correct responses (a better definition for RR would be Success Rate, but we use RR to be consistent with the literature), whereas RR_m considers both correct and incorrect response and their relative weight. Furthermore, RR_m is the optimal strategy in paradigm with fixed task duration (see Materials and Methods; e.g., Simen and Cohen, 2009; Bogacz et al., 2010). Harris et al. (2014) have hypothesized that the RR_m may also account for the observations that response times have a reciprocal normal distribution (Carpenter, 1981). As it will turn out, the distributions resulting from the EXACT paradigm, with the current design, does not appear to be reciprocally normally distributed, so this point will be considered more deeply in the discussion.

The purpose of this study was threefold. First, we were interested to see how participants responded to this novel paradigm. In particular, we explored whether participants would choose a response time RT and a response gauge position $ACC(RT)$ that depended systematically on our manipulation of the $ACC(t)$ parameters α and λ. Second, we asked whether the data could allow us to distinguish among the four decision rule models. Third, we use the results from the paradigm to explain patterns of behavior that have usually been explained with a rise-to-threshold accumulator process.

MATERIALS AND METHODS

At the beginning of each trial, a gauge was shown on a monitor screen (**Figure 1**). The gauge started at a predetermined level and "filled up" according to $ACC(t)$ (Equation 1). The participant pressed the CTRL key to request a reward. Reward

and punishment were in form of *points*. When the participant requested a reward, 1 point was awarded with probability given by the level of the gauge (i.e., $ACC\,(t)$), or 1 point was taken away with probability $1 - ACC(t)$ (punishment/reward ratio $q = 1$). After every response, a text saying "Please Wait" appeared for 0.5 s (delay d). When a reward was awarded, "+1" in green appeared on the screen, and "–1" in red appeared for a loss (**Figure 1**). Each participant started with 25 points at the beginning of each condition. During each condition, the screen also displayed how many minutes were left as well as the current total accrued points. Each condition lasted 3 minutes regardless of performance (no participant lost all their points), so that faster responses would lead to more trials (and possibly more profit). Participants were instructed to win as many points as possible. They were informed on the relationship between the gauge level and the probability of getting/losing a point, and were given a practice session.

This study was approved by the Faculty of Health and Human Sciences Ethics Committee at Plymouth University, with written informed consent from all participants. All participants gave written informed consent in accordance with the Declaration of Helsinki.

Procedure

We recruited 17 female and 3 male Psychology students to take part in the study. To increase motivation, they were informed that the participant who won the most points would receive a £10 Amazon voucher. Each participant underwent 2 × 4 conditions. The starting point of the gauge was set to either $\alpha = 0.25,$ or 0.75, and the speed of the gauge was set to $\lambda = 0.166,$ 0.33, 1, or 2. This means that at the very beginning of the trial the chance of receiving a reward was equal to α and the chance of receiving a punishment was equal to $1 - \alpha$. After that, the chance of receiving a reward increased according to the $ACC(t)$ (the chance of getting a punishment decreased accordingly). Higher values of λ indicates faster increase. Sequences were randomized across participants, except that each participant was tested on all the speed conditions for a particular starting point, before being tested on the other starting point condition. The participants performed an initial training session to make sure they understood the task ($\alpha = 0.5$, $\lambda = 1$). They were then given another training session at the beginning of each starting point condition (in the training we set α equal to the starting point for that condition and $\lambda = 1$). They were allowed to take a break between conditions. There was no response deadline.

Decision Rules Predictions

Each decision rule generates different predictions based on the assumed $ACC(t)$. The mathematical formulation of the four decision rules is similar to the one in Bogacz et al. (2006), with few minor differences: firstly, they use the parameter "Error Rate" ($ER(t)$) instead of $ACC(t)$. The parameterization in terms of $ACC(t)$ was found to be more convenient for the experimental design employed here ($ER\,(t) = 1 - ACC(t)$). Conversely to Bogacz et al., the additional penalty delay following an error was ignored, since it is not used in our experiment and it is not common in classic RT paradigms. We also did not separate decision time (the time in which the information is accumulated and therefore accuracy increases) from non-decision time (sensory and motor processing), because in the EXACT paradigm any time spent on the trial increases accuracy.

TABLE 1 | Summary of decision rules with associated optimum response time t^* for the exponential $ACC(t)$ in Equation (1).

Decision rule	Optimum response (t^*) by assuming $ACC\,(t) = 1 + (\alpha - 1)\,exp\,(-\lambda t)$
$BR = -\left[t + q\,(1 - ACC\,(t))\right]$	$t^*_{BR} = -\dfrac{\ln\left(\dfrac{1}{\lambda q(1-\alpha)}\right)}{\lambda}$
$RR = ACC(t)/(t + d)$	$t^*_{RR} = -d - \dfrac{W_{-1}\left(\dfrac{\exp(-\lambda d - 1)}{\alpha - 1}\right) + 1}{\lambda}$
$RA = RR - \dfrac{q}{d}\,(1 - ACC\,(t))$	No explicit form
$RR_m = \dfrac{ACC\,(t) - q\,(1 - ACC\,(t))}{t + d}$	$t^*_{RR_m} = -d - \dfrac{W_{-1}\left(\dfrac{\exp(-\lambda d - 1)}{(\alpha - 1)(q + 1)}\right) + 1}{\lambda}$

The mathematical formulation of the decision rules with the corresponding optimum response time t^* is shown in **Table 1**. In these formulae d is the delay across trials, t is the time from the beginning of the trial to the participant's response, q is the weight of accuracy relative to the speed of reward (Bogacz et al., 2006) and it is assumed to be subjective and dependent on individual differences.

We derived the optimum time of responding t^* (right column of **Table 1**) by setting the derivative of the decision rule to 0 and solve for t (see Appendix in Supplementary Materials for details). **Figure 2** shows both the optimum t^* as a function of λ for different parameters value (how long a participant should wait before responding, left panel), and the $ACC(t^*)$ value (what accuracy level is reached upon responding, right panel), assuming the exponential $ACC(t)$ in Equation (1). For most models, the optimum response t^* goes to 0 when λ goes to 0, which means that when the accumulation of evidence in time is too slow (because the trial is difficult, the stimulus is dim, etc. . .), then it is not convenient to spend long time accumulating information. This is not true for the RR_m model with some combination of parameters α and q. In particular, the optimum t^* goes to infinity when λ goes to 0 given that $\alpha < \frac{q}{q+1}$ (see **Figure 2**, bottom panels, and Appendix in Supplementary Materials for derivation) and goes to 0 when $\alpha < \frac{q}{q+1}$. The RR_m model is particularly relevant because it is the only criterion that actually maximizes the gain for a task (including but not limited to the EXACT paradigm) with a fixed task duration (in subjective utility value, see Harris et al., 2014)[2]. For RR_m the parameter q can be interpreted as the punishment/reward ratio. The experimenter can try to affect this parameter (by using

[2]To show why this is true, note that $c_1 ACC\,(t) - c_2\,(1 - ACC\,(t))$ (the numerator in RR_m) is the average gain of a subject for a trial, where c_1 and c_2 are the subjective utility (reward and punishment). For an experiment lasting l min, the subject will be able to perform $n = l/(t + d)$ trials, where t is the response time of the subject, d the other delay of the trial. Therefore to get as many points as possible for an experiment, the participants should maximize $[c_1 ACC\,(t) - c_2\,(1 - ACC\,(t))]\ \frac{l}{t+d}$, which maximum correspond exactly to $t^*_{RR_m}$, with $q = c_2/c_1$.

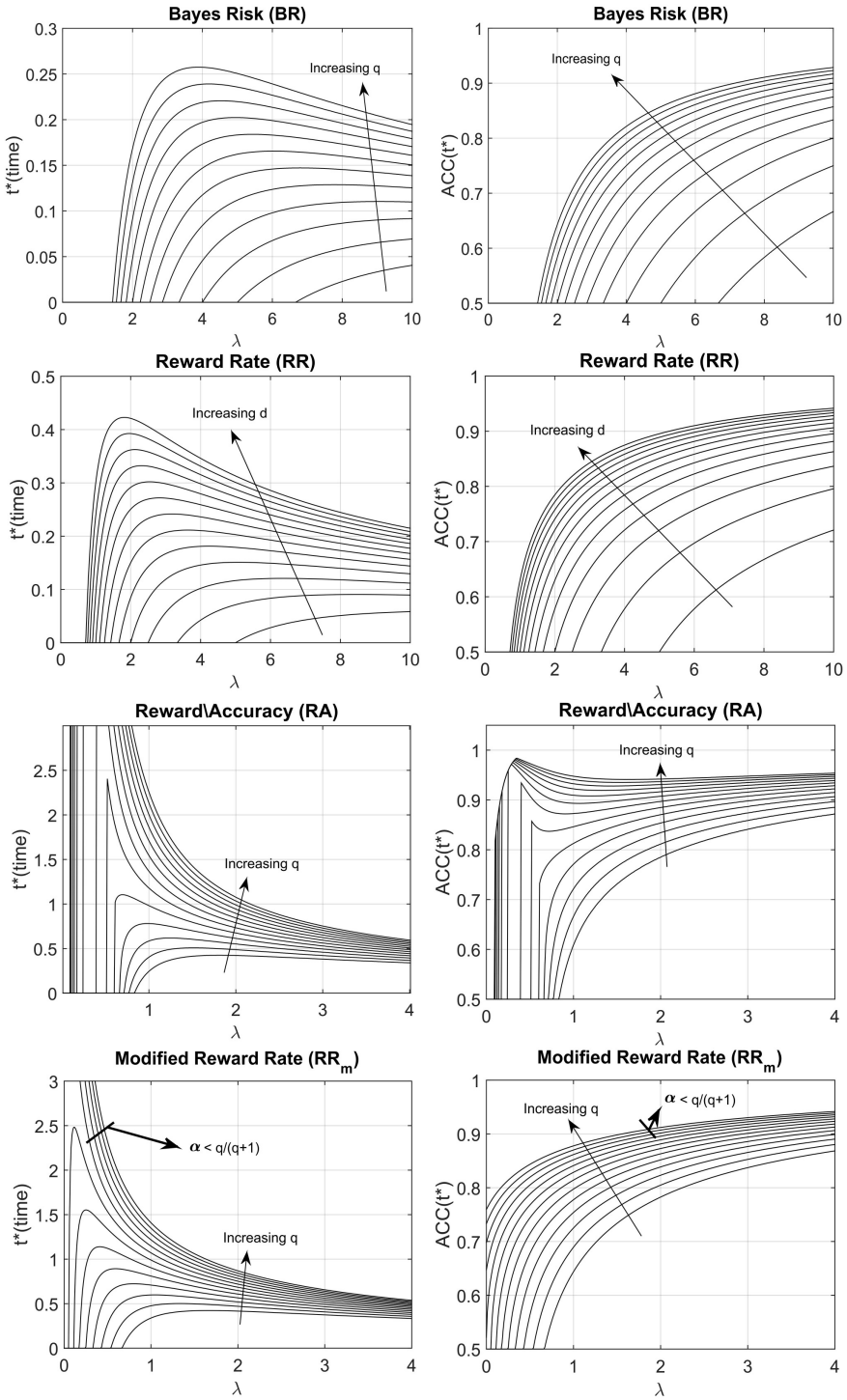


FIGURE 2 | Optimal response time, *t (left column) and ACC(*t**) (right column) vs. λ for the four decision rules, with different parameters q and d.** *ACC*(*t**) indicates the optimum level of accuracy that has to be reached upon responding. The arrows indicate how the shape changes when the d or q parameter is increased. The q parameter represents the subjective punishment/reward ratio; the d parameter represents the length of the trial, excluding the reaction time. Increasing q or increasing d corresponds to an increase in the optimum response time (*t**) and a consequent increase of ACC(*t**). For the Modified Reward Rate (**bottom panels**), two very different *t** shapes are produced with different combinations of q and α (the starting point of the gauge).

monetary reward, sound feedback, instruction set, etc…), but other characteristics may be more relevant for the participants, such as intrinsic loss avoidance.

Statistical Analysis

Due to small sample size (sometime less than 15 responses) obtained in some conditions (especially with small λ), we aggregated the data by calculating the median response for each participant. We fitted the three decision rules (BR, RA, RR_m) response predictions (the condition-dependent t^* shown on the right side of **Table 1**) to the aggregated median responses by an optimization procedure that minimizes the least squares (MATLAB *fminsearch*). RR did not need to be fitted because it is parameter free (d is assumed to be known by the participant to be equal to 0.5). For BR, RA and RR_m, the parameter q was estimated with two different approaches: by leaving q free to vary across the two α conditions or by fixing it to be equal for the two α conditions. The value of q was bounded to be positive.

To analyse how RT and $ACC(RT)$ changed within a particular condition, the responses were binned for each participant and each condition according to the time elapsed from the beginning of the trial. This was made in order to have the same dataset size for each participant, regardless of the number of responses produced. For the $\lambda = 0.166$ and $\lambda = 0.33$, bins of 10 s were used and for $\lambda = 1$ and $\lambda = 2$, bins of 5 s were used. The difference in bins was due to generally slower RT with smaller λ in the $\alpha = 0.25$ condition, and longer bins were needed to include enough datapoints for reliable statistics. The mean and standard deviation were calculated for each bin, resulting in two time series for each participant and each condition. We computed two linear regressions (mean against time and standard deviation against time) for each dataset, and the resulting slopes were analyzed with a Two-way repeated measure ANOVA and a multiple comparison test.

To understand the sequential dependency along trials, the average partial autocorrelation function (PACF) for each participant and each condition was computed. This function returns the autocorrelation between a response RT_t and RT_{t+lag} removing the dependency through RT_{t+1} and $RT_{t+lag-1}$. We used a maximum lag of 10. The average PACF across all participants was computed by averaging the individual PACF values for each condition.

Participants' RTs were grouped in order to have a single distribution for each condition by using the Vincentizing method (Ratcliff, 1979). To calculate the response in the rate domain, we computed the rate for each response (1/RT), then standardized the data into z-scores for each participant and each condition, which allowed us to collapse all the data together (Harris et al., 2014).

Any $ACC(RT)$ distribution must be truncated because $ACC(RT)$ is bound to lie between 0.25 or 0.75 (depending on the α condition) and 1. We wanted to check whether the $ACC(RT)$ distribution generated by each participant was close to a truncated Normal. This was complicated by small individual samples, and by the fact that aggregating truncated Normal with different moments makes it ambiguous whether the original distributions were truncated Normal or not. We excluded the conditions where truncation was severe, that is when the sum of mean and two standard deviations was over one and the difference between mean and two standard deviations below α (the truncation point). The value of two for the standard deviation was chosen as a balance between excluding truncated datasets and not excluding too many data points (see Harris et al., 2014). The remaining distributions were standardized into z-scores and collapsed together. By excluding the most severely truncated distribution and aggregating the remaining standardized distributions, we expect to obtain a normal distribution with a milder truncation, if the original distributions were truncated Normal.

RESULTS

The examination of the total score gained (points at the end of each condition) can be used to understand how participants responded to the task. Their score can be compared to the optimum average score (expected amount of point at the end of each condition, given that participants were actually trying to maximize the amount of point earned). The optimum expected score was computed by finding the maximum points that could be won for a fixed task duration (RR_m with $q = 1$ and $d = 0.5$) (**Figure 3**, solid circles). For $\alpha = 0.25$, the total score increased systematically with λ (**Figure 3A**, open circles), and clearly demonstrates that participants were sensitive to the speed of the gauge. As the gauge level increased it was possible to obtain more points in the available 3 min, and although there was variability among participants, some were close to the optimum performance.

For $\alpha = 0.75$, the optimum response is $t^* = 0$ for all values of λ, and the maximum score was always 205 points (filled circle in **Figure 3B**). Participants scores were more variable than before and depended on λ, with an increasing mean and decreasing variance with λ. Even though participants could not possibly reach a $RT = 0$, they should have had the same average RT for the four λ conditions when $\alpha = 0.75$, if they were maximizing RR_m with $q = 1$ and $d = 0.5$.

Figure 4 shows responses for one representative participant in terms of RT (blue lines) and positional responses along the gauge, $ACC(RT)$ (orange lines). The positional response $ACC(RT)$ were variable but typically showed an increasing trend with higher λ. Positional responses were lower for $\alpha = 0.25$ than 0.75. This participant was selected because his/her responses followed quite closely the trend found in aggregated data for all the conditions.

Note that high variability in RT does not always correspond to high variability in $ACC(RT)$ because RT depends on on the speed of the gauge [see Equation (2)]. For example, in **Figure 4**, when $\lambda = 0.16$ and $\alpha = 0.75$, RTs vary widely across trials, but corresponds to only a slight variation in $ACC(RT)$.

We next examined median response times mdRT and median positional response $ACC(mdRT)$, across all (aggregated) participants for each condition (**Figures 5A,B**). For each participant and condition, the first 10 responses were omitted to avoid contamination from any potential adaptive/learning at the start of a condition. There was a significant effect of both

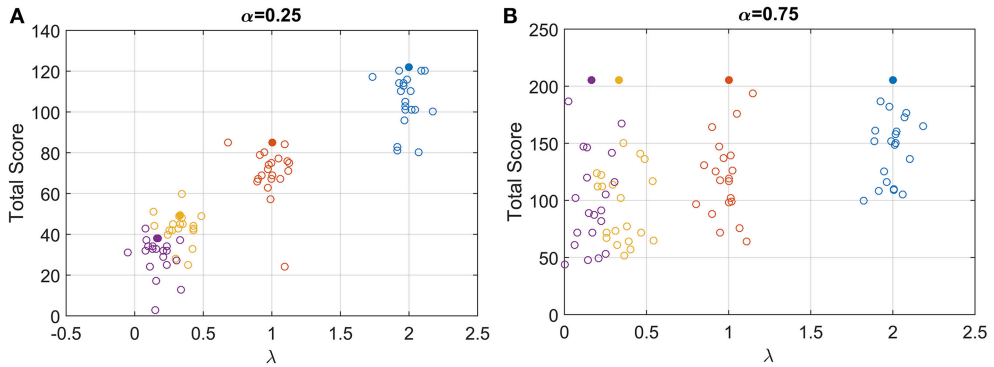


FIGURE 3 | Open circles: total score for each participant for the low (A) and high (B) starting point condition. For clarity, each circle is plotted with a random horizontal offset. **Filled circles:** expected score with optimum performance for the four speed conditions and the two starting points conditions.

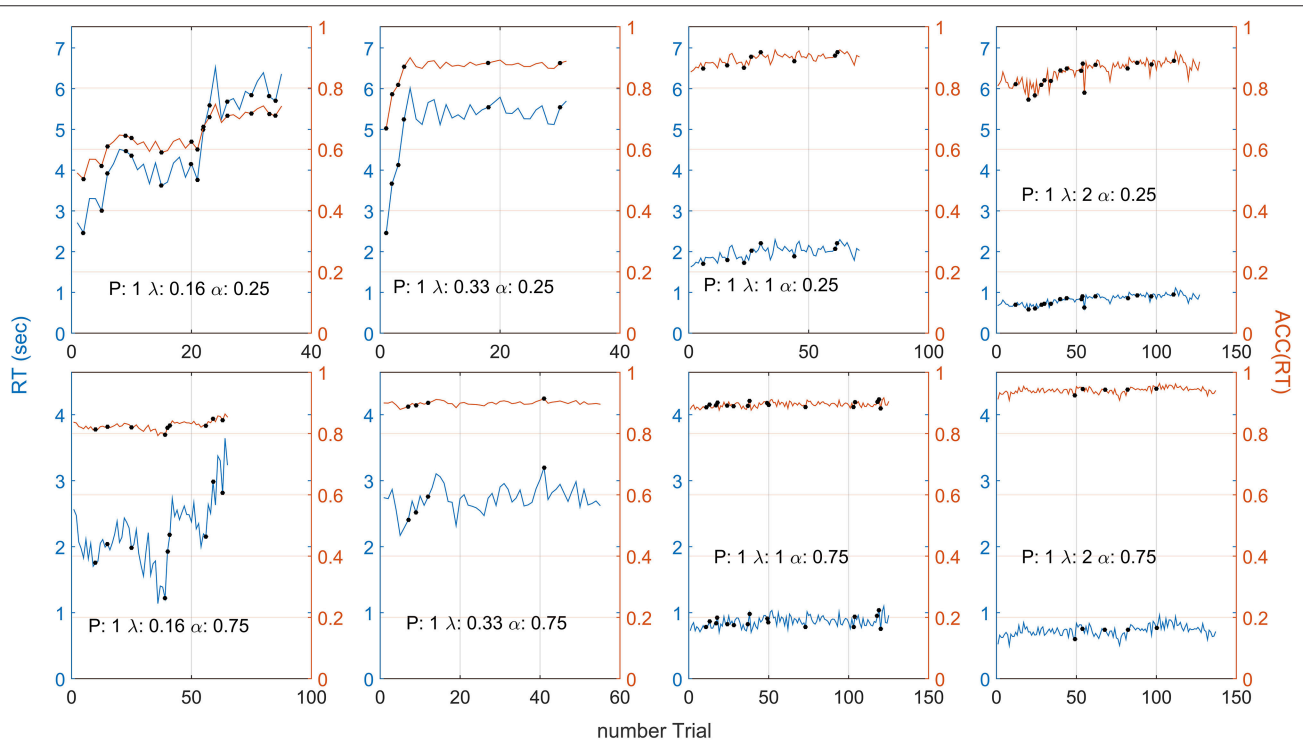


FIGURE 4 | Participant 1 responses in terms of RT (blue lines) and *ACC*(*RT*) (orange lines) for each λ and α condition. The black dot indicates that a punishment has been given on that trial. On average, *ACC*(*RT*) increases with λ and is usually lower for the $\alpha = 0.25$ condition. Note that the different number of trials depended on the condition (which always lasted 3 minutes). For example, when $\alpha = 0.25$, $\lambda = 0.16$, the gauge grew more slowly, participant's RT were slower, and less trials were made. Note that high variability in RT does not always correspond to high variability in *ACC*(*RT*). Compare this plot with **Figures 5A,B**, for the grouped RT and *ACC*(*RT*) across participants. Note that the two α conditions have two different scales.

the starting point α and speed λ on mdRT (Two-way repeated measures ANOVA; for λ: $F_{(3,\ 152)} = 24.61$, $p < 10^{-12}$; for α: $F_{(1,\ 152)} = 53.2$, $p < 10^{-10}$, and for the interaction $\alpha \times \lambda$: $F_{(3,\ 152)} = 10.84$, $p < 10^{-5}$).

For $\alpha = 0.25$, mdRT decreased monotonically with λ. We were able to capture these data by a power law function $mdRT_i = \gamma + k\lambda_i^{-\beta}$ for $i = 1, 4$ (yellow line in **Figure 5A**). The estimated values were $\gamma = 203ms$, $k = 1.45$, $\beta = 0.80$, with $R^2 = 0.997$, indicating an excellent fit. This implies an equivalence between λ and conventional stimulus intensity in classic Piéron's law (see Discussion). For $\alpha = 0.75$, mdRT was much less than for $= 0.25$, and did not follow a power law. Moreover, mdRT was not monotonic and exhibited a small peak at $\lambda = 0.33$.

Similarly *ACC*(*mdRT*) clearly depended on α and λ. This means that participants were actually adjusting their positional response for different conditions and not responding to a fixed position along the gauge independently of λ and α.

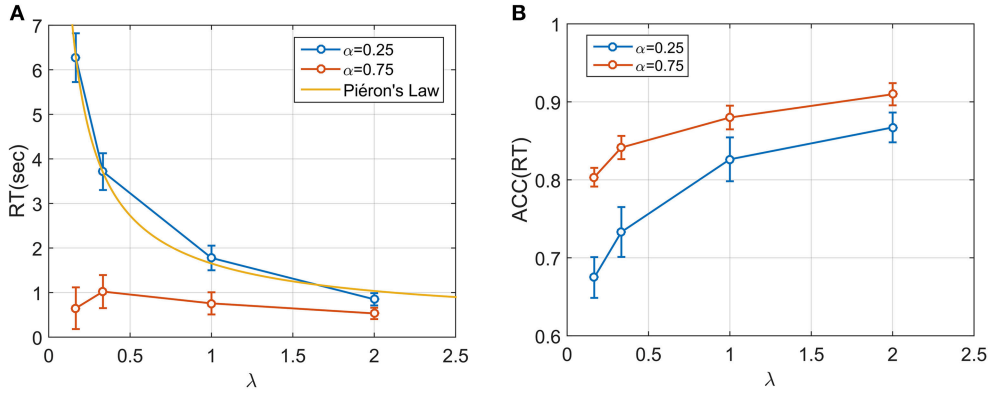


FIGURE 5 | (A) Median RT across the 20 participants for four speed conditions and two starting point conditions. A power (Piéron) function was fitted for the low starting point condition ($\alpha = 0.25$). Median RT appear to be dependent both on λ and α and on their interaction. In particular, when $\alpha = 0.75$ the mdRT were not monotonic with a peak at $\lambda = 0.33$. **(B)** *ACC*(*mdRT*), showing how participant adjusted their positional response along the gauge based on both α and λ.

Fitting the Models to the Aggregate Data

The mdRT were fitted with the optimum t^* predictions from four decision rules: Bayes Risk, Reward Rate, Reward/Accuracy, Modified Reward Rate (**Figure 6**). We minimized the residual least squares by allowing q to vary and to be fixed across the two α conditions (this does not apply to RR which does not have free parameters). The plots when q is allowed to vary are shown in **Figure 6** and the residual least squares in **Table 2**. When q was allowed to vary across α conditions, the only model that could provide a reasonable fit was RR_m. The estimated q (showed in **Figure 6**) clearly differed from the two α conditions (see Discussion). When the parameter q was fixed the resulting curves did not seem to be able to capture the different shape between the two α conditions. In particular the residuals for $\alpha = 0.25$ were similar for the two fitting methods, but they were higher with fixed q in the $\alpha = 0.75$ condition (see **Table 2**). Note that with $\alpha = 0.75$, a residual value of 2.32 corresponds to a horizontal zero line.

Time Series Analysis

We examined sequential dependencies of RT by calculating the partial autocorrelation function (PACF) averaged across participants and conditions (see Materials and Methods). The mean PACF with lag one was 0.457, and with lag two was 0.090. Higher order lags were negligible. Thus, responses depended explicitly on the previous two responses. To explore this further, we examined how this sequential dependency depended on the reward/punishment on previous trials. We computed differences between consecutive RT responses depending on whether the previous response had been punished or rewarded. These differences were averaged across participants and conditions. For a punishment, the mean change in RT was an increase of 345.92 ms (SE: 97.95) and for a reward, a decrease of 61.88 ms (SE: 30.32 ms). For two punishments in a row, the mean increase was 583 ms (SE: 139.26), and for two rewards in a row, the mean decrease was 94.01 ms (SE: 28.62). For three or more rewards/punishments in a row, there was little additional change, consistent with the PACF results. Repeating the analysis for *ACC*(*RT*) showed a similar pattern. After one punishment there was a mean increase of 0.0218 (SE: 0.0058), and after two punishments in a row, a mean increase of 0.048 (SE: 0.0082). After a reward there was a mean decrease 0.0038 (SE: 0.0022), and after two rewards a decrease of 0.0066 (SE: 0.0020). We also analyzed how the 1st order changes depended on λ and α. The size of RT adjustment after a punishment became larger for smaller λ, but with no obvious dependency on α (**Figure 7A**). However, the change in *ACC*(*RT*) showed the opposite pattern, with strong dependency on α and a weaker dependency on λ (**Figure 7B**).

The time course of participants' responses during a condition showed variability and small trends. The mean and standard deviation for each participant's response during a condition was computed by dividing the total duration (3 minutes) into bins of 5 or 10 s depending on the condition (see Materials and Methods), and then we performed a linear regression against time. There was little difference between RT and *ACC*(*RT*), so here we consider only RTs. Considering the changing in RT mean across condition duration, we found that the slope was significantly affected by both λ and α conditions (Two-way repeated measure ANOVA; for λ:$F_{(3,\ 152)} = 10.49$, $p < 10^{-13}$, for α:$F_{(1,\ 152)} = 12.15$, $p < 0.001$). However multiple comparison tests on the slopes of the fitted regression model showed that the only significantly different condition (at 0.01 critical level) was found for $\alpha = 0.25$ and $\lambda = 0.16$, which corresponded to an average increase of 13.5 ms for second (99% CI [9 ms, 17 ms]). None of the other slopes were significantly different from zero. An analysis on the slope of standard deviation of RT vs. time revealed no significant differences from zero for any of the λ and α conditions.

Distribution Analysis

The resulting distributions collapsed across participants are shown in **Figure 8**. The distributions are right skewed (apart from the condition $\lambda = 0.16, \alpha = 0.25$). We used the estimated q values found with the aggregated data (shown in **Figure 6**) to plot for each condition the four decision rules functions (colored

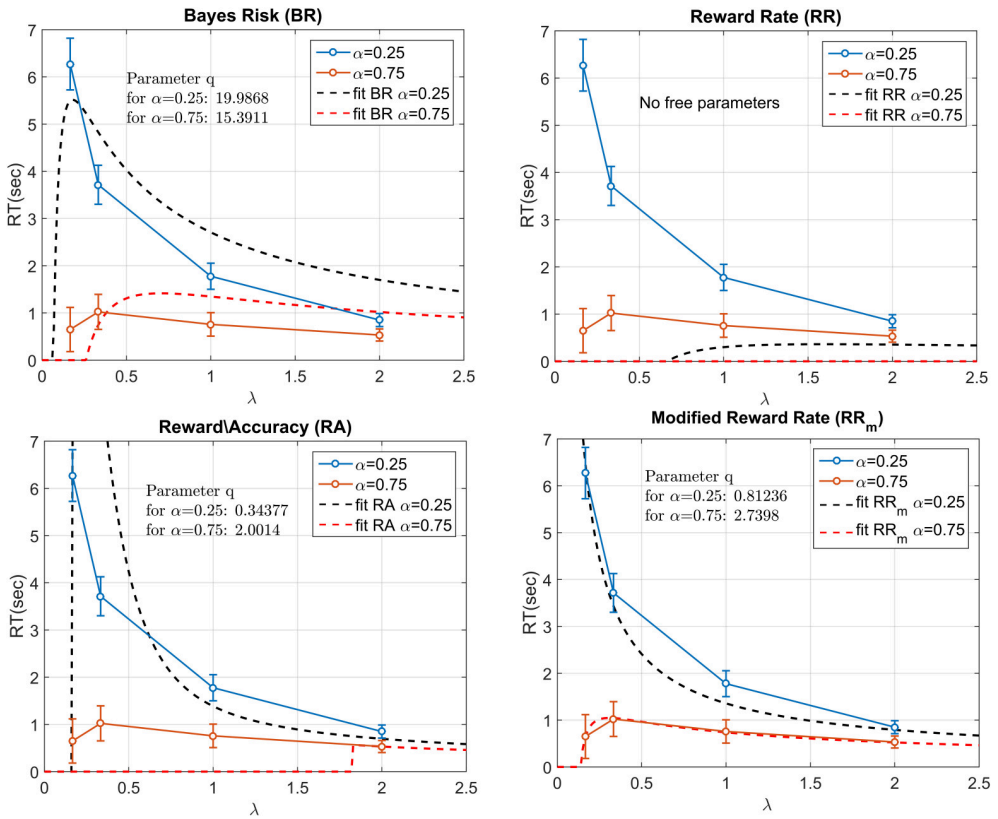


FIGURE 6 | Fitting the four decision rules to the data by minimizing the least squares. RR does not have any free parameter, whereas for the other decision rules the free parameter q had to be estimated. In the figure, the parameter *q* is allowed to vary across the two α conditions. The best fit was provided by RR_m. The two estimated q values were largely different between the two α conditions.

TABLE 2 | Residual least squares for each decision rule and the two starting point conditions.

	$\alpha = 0.25$		$\alpha = 0.75$	
	q free	**q fixed**	**q free**	**q fixed**
BR	3.43	3.53	1.08	1.55
RR	55.56		2.32	
RA	36.64	36.64	2.04	2.32
RR_m	0.29	0.29	0.0009	2.32

*Models were fitted by fixing q across α conditions or by allowing it to vary (this does not apply to RR, which does not have any free parameter). A residual least square value of 2.32 for $\alpha = 0.75$ corresponds to a horizontal zero line. Fixing q resulted in consistently poor fit with the $\alpha = 0.75$ conditions. The best fitting decision rule was RR_m. When q was allowed to vary, RR_m provided an excellent fit of empirical data for both the α conditions (see **Figure 6**).*

curves), representing the expected gain for that decision rule (the equation for each decision rule function is shown in **Table 1**, left column). If a participant is trying to maximize a decision rule, she should always respond precisely at t^* (filled circles in **Figure 8**). With time estimation error, temporal uncertainty, and other source of sensory/motor noise, participants might respond according to a distribution that roughly follows the shape of the decision rule function, with the maximum point slightly ahead of the distribution mean (see Section Asymmetricity of the decision rule and optimization algorithm for further discussion). In our data RR_m is the only decision rule that follows this pattern (the only exception being for $\alpha = 0.75$ and $\lambda = 0.16$).

The distribution in the rate domain was analyzed by using the standardization approach (see Materials and Methods). The distributions did not appear normal as observed in choice RT, but were positively skewed. **Figure 9** shows the standardized rate distributions collapsed across all the participants and conditions (there was no significant difference between different conditions).

We also examined the distributions of $ACC(RT)$. Because of the possible left truncation at 0.25 and 0.75 for the two α conditions, and right truncation at 1 for both conditions, we excluded severely truncated dataset (see Materials and Methods). The remaining 89 out of 160 (20 participants $\times 2 \times 4$ conditions) mildly truncated datasets were standardized and collapsed across subjects. **Figure 10** shows the original, standardized and aggregated distributions (in blue) and the modified standardized and aggregated distributions (in black). As a comparison, a normal distribution is fitted to the modified dataset. Note that for most of the conditions, excluding the truncated datasets does not have a big effect, and both distribution pre- and post-exclusion are approximately truncated Normal. The two exceptions are the distributions in the $\alpha = 0.75$ and $\lambda = 0.16$, 0.33 conditions,

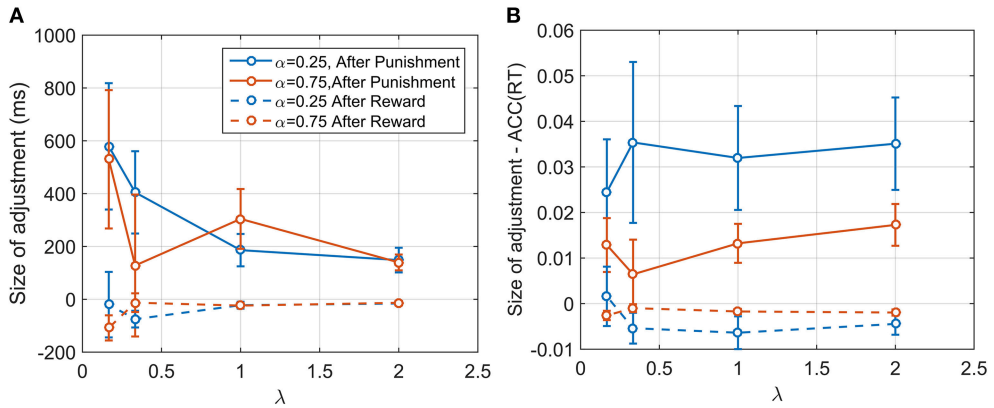


FIGURE 7 | (A) Size of the adjustment in ms after a punishment and after a reward for α and λ.The adjustment is larger after a punishment, and seems to be dependent on the conditions itself. **(B)** Size of adjustment in terms of *ACC*(*RT*). As before, the adjustment is larger for punishment. Contrary to what shown in **(A)**, the plot reveals that the size of adjustment is dependent on α and more weakly on λ. It is plausible that participants were adjusting their accuracy by a certain proportion, and in so doing changed RT depending on λ.

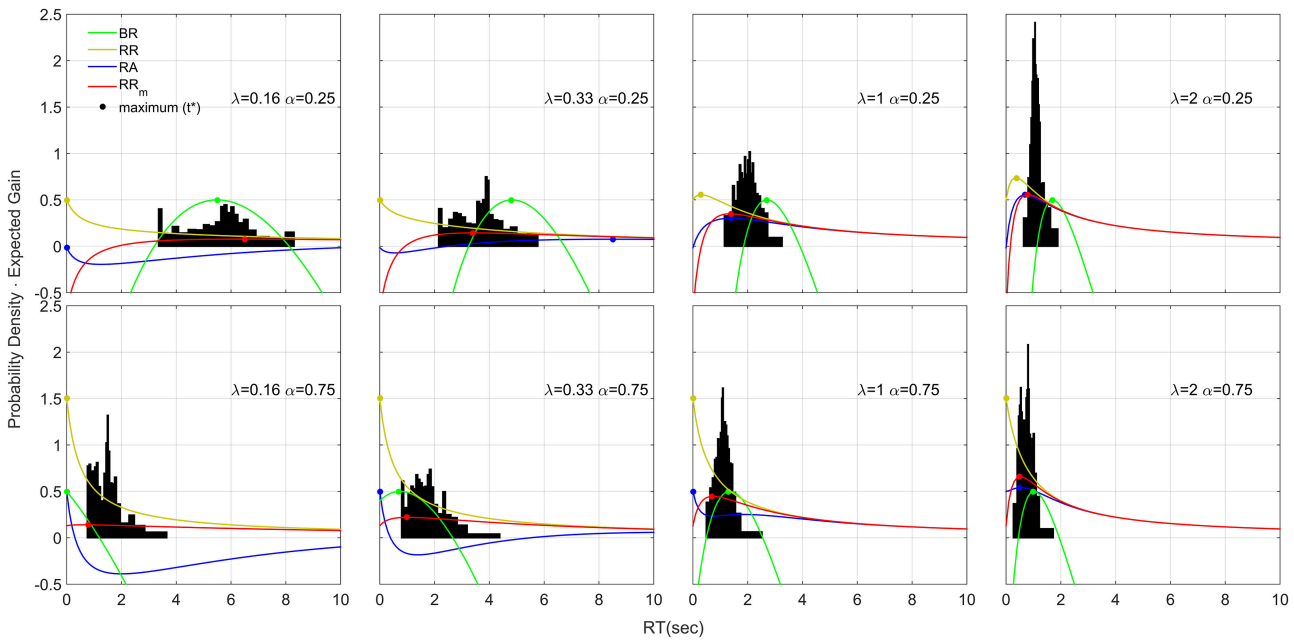


FIGURE 8 | Response Distributions for each condition obtained by averaging the quantiles across each participant and each condition (Vincentizing). The distributions are scaled so that the area of each one is unity. The distributions appear to be skewed on the right side and their shape depends on λ and α. The colored lines refer to four decision rule functions. Note that with respect to the decision rules, the vertical axis corresponds to the expected gain given that particular decision rule. To increase visibility we shift the BR function up by subtracting to the function its maximum value (which was negative) and adding 0.5. In this way the maximum of the BR is always 0.5, but the shape is unchanged. Note that the absolute value of the function is not important for the distribution shape, but the time-depending shape of the decision rule function is. Therefore, shifting the function up or down does not have any theoretical implication for the distribution shape.

which were skewed and clearly not normal before processing. These two conditions also contained the highest number of severely truncated distributions (only 7 and 8 datasets remaining, see **Figure 8**). These results are in accord with the hypothesis that accuracy of responses is normally distributed and with most severe truncation with when λ is small and α is high (see Section Accuracy is normally distributed).

DISCUSSION

Participants' responses to the trials clearly indicated sensitivity to the paradigm. Participants did seem to be attempting to win points, as can be seen in **Figure 3**. Their median response positions along the gauge depended systematically on the accuracy function $ACC(t)$ (**Figure 5A**). When the gauge level

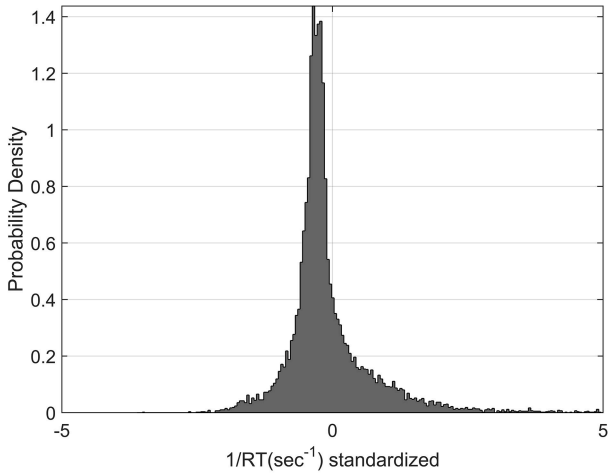


FIGURE 9 | Standardized Rate Distributions obtained by collapsing the standardized rate for all participants and all conditions (see Materials and Methods). The distribution is clearly not normal. Similar shapes were found when analysing each condition separately.

moved faster (higher λ), participants' RT decreased, and they selected a position further along the gauge [increasing $ACC(RT)$]. This implies that participants were not simply choosing some fixed or idiosyncratic position for all λ and α conditions, which would correspond to a decreasing RT but equal $ACC(RT)$, but they were actually changing strategy for each λ condition.

Their response also depended on the starting point. When α was lower, participants on average waited longer, and their accuracy decreased. The dependence on λ was strongly dependent on the starting point. Participants slowed their response when $\alpha = 0.25$ and $\lambda = 0.16$ (very slow gauge speed). However, the same did not occur with $\alpha = 0.75$ and a slow gauge speed. In this condition, most of the participants realized that slowing their response would not correspond to a significant increase of accuracy, and therefore their response was faster.

Comparison of the aggregated data to models supported the RR_m decision rule. The fitting with RR_m, when the parameter q was allowed to vary, was better than the other models (**Figure 6**). For all the models, included RR_m, fixing q resulted in a worse fit. The q parameter, when compared with experimental reward and punishment, can be interpreted as a measure of risk attitude. Thus, risk aversion/seeking corresponds to observed q being greater/lower than actual punishment/reward ratio (1 in our experiment). In our experiment, estimated values of q were quite dissimilar for the two starting point conditions, with significantly higher q for the $a = 0.75$ condition. This suggests that participants changed their subjective punishment/reward ratio and become more cautious with increasing α. This is consistent with the "fourfold pattern of risk attitudes" in prospect theory (Tversky and Kahneman, 1992), according to which risk-averse behavior (in our terminology, setting a high q) are more common when gains have high probability (like in our $\alpha = 0.75$ condition). Balci et al. (2011) found that the estimated q values for the RR_m decision rule decreased with training, to ∼0.2. The short session duration in our experiment did not allow us to perform the same analysis, but is in principle possible to with the EXACT paradigm.

There may also be a gender difference in the decision rules in the distribution of parameters. Bogacz et al. (2010) found a clear difference in male and female performance, female responses being less optimal (more risk averse) with respect to the RR strategy. The EXACT paradigm could be similarly (and, we believe, more conveniently) used to investigate whether difference in attitude across gender may be captured by a different estimation of the q parameter and hence risk aversion. Unfortunately, this analysis is not possible in the current work due to the severe imbalance between female and male, which reflects the female-male ratio in Psychology classes in our University.

Even though RR_m is the decision rule that maximizes the rate of gain for any task with a fixed duration (Harris et al., 2014), it has been found (Bogacz et al., 2010) that for some participants, RA may be a better model than RR_m. However, Bogacz et al. assumed the drift diffusion model. This assumption constrained the set of predictions generated by the decision rule. For some participants the drift diffusion model may not be a good model of the perceptual process, which may affect the goodness of fit of the decision rule. By using the EXACT paradigm, the current work overcomes this problem.

The EXACT paradigm can be considered as an abstract version of a classic RT task, in which more factors are controlled (more specifically, all the accumulation usually performed by the perceptual process is artificially manipulated by the researcher). We believe it is important to compare this paradigm with the classic RT design. Disentangling the perceptual process from the decision rules mean that is possible to understand which features of the experimental data are due to one or the other part of the decision mechanism. For example, the right-skewness of the distributions that we observed in our dataset is usually deemed to be the result of a rise-to-threshold diffusion process. There is no reason why they should emerge in our paradigm because $ACC(t)$ is exogenous. However, this does not mean that assuming a rise-to-threshold mechanism is unnecessary in a classic RT task. The $ACC(t)$ in a perceptual task has to emerge from a process that accumulates information, and a rise-to-threshold model seems plausible. Nevertheless, these distribution features can be explained by an alternative process implying that the rise-to-threshold model may be modified/re-interpreted differently.

The results along conditions for an individual participant (**Figure 4**) and the size of adjustment after a punishment or a reward (**Figure 7**) show an important point about the EXACT paradigm: the possibility to analyse the result not only in terms of RT, but also in terms of $ACC(RT)$. This is not possible in the classic RT paradigm because $ACC(RT)$ is not known by the experimenter, unless a particular perceptual process is assumed. The analysis of $ACC(RT)$ gives additional information about the strategy employed by the participants. For example, by analysing the size of adjustment in terms of RT after a punishment/reward, we would have concluded that participants adjusted their response based on the speed of the gauge itself. However analysing the data in terms of $ACC(RT)$ reveals that participants may be focusing on adjusting the positional response

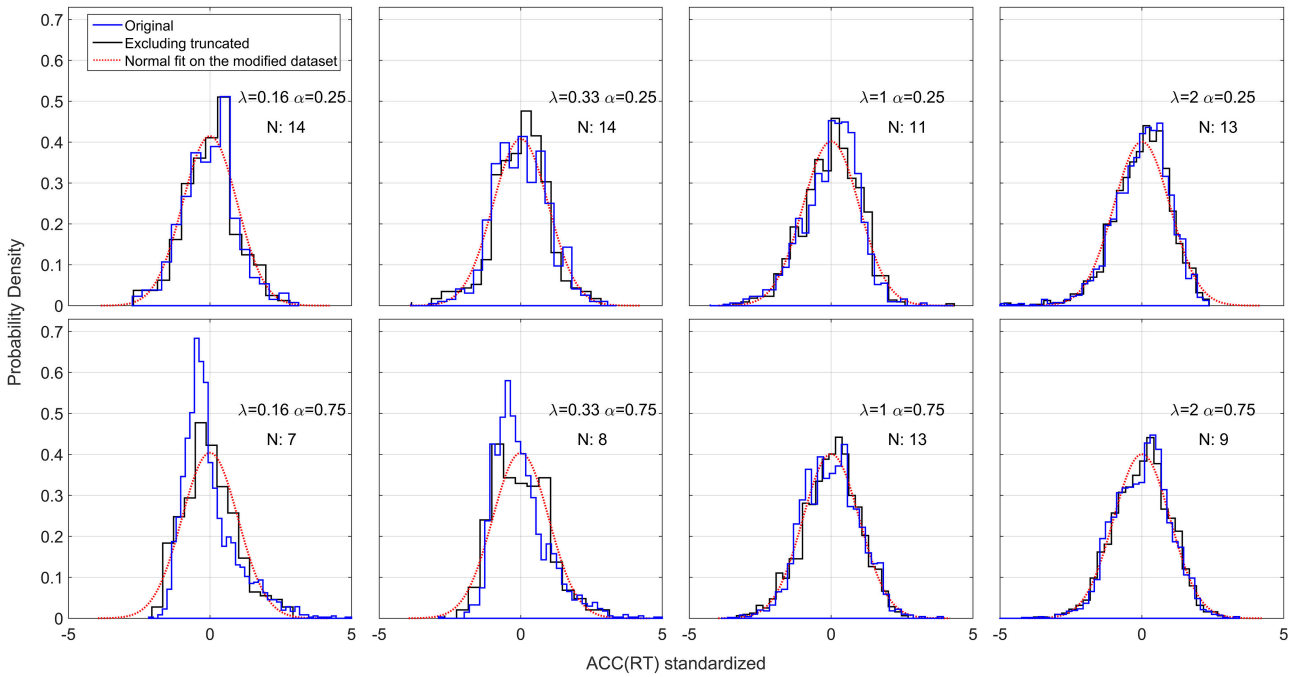


FIGURE 10 | Standardized positional response along the gauge: *ACC(RT)*. The blue lines represent the original standardized and collapsed distributions; the black lines represent the standardized and collapsed distributions after eliminating the datasets with severe truncation. *N* represents the number of remaining datasets (out of a total of 20 participants) after eliminating severely truncated distributions. A normal density is fitted to each modified distribution.

along the gauge, trying to increase or decrease their accuracy at the same proportion for different λ.

The EXACT paradigm has some similarities with the "Beat the Clock" task (Simen et al., 2011), in which a participant earns a reward that is an exponential function of time. If the participant responds after the deadline the reward is zero. In our task, waiting increases the probability of getting a reward, but not the reward itself. Furthermore, in our paradigm participants know perfectly the function shape, as it is shown on the monitor screen, whereas in the "Beat the Clock" they have to infer it. Finally, Simen et al. did not fix the condition's length, which is part of our paradigm. Generally, "Beat the Clock" seems more suited for investigating temporal discrimination, whereas the EXACT paradigm is more suited for analysing decision rules.

Piéron's Law

We found that a power law provided an excellent fit for the mdRT in the $\alpha = 0.25$ condition: $mdRT_i = \gamma + k\lambda_i^{-\beta}$. In psychophysics, a similar power law describes the relationship between stimulus intensity and mean RT, (Piéron's Law). During the last century, the law has been reported to match observed data in numerous sensory modalities and experimental paradigms (e.g., brightness detection: Piéron, 1914; color saturation: Stafford et al., 2011; taste signals: Bonnet et al., 1999; simple and choice reaction times: Pins and Bonnet, 1996). Obtaining a Piéron's function with the EXACT paradigm strongly supports the idea that λ corresponds to stimulus intensity in a classic RT paradigm. We estimated these parameters to be $\gamma = 203$ ms, $\beta = 0.8$, and $k = 1.45$. These values are consistent with the literature for Piéron's law (Luce, 1986). Within this new framework, β can be interpreted as being dependent on the relationship between λ and the physical stimulus intensity.

Recently, it has been proposed that Piéron's Law is a necessary consequence of rise-to-threshold decision making (Stafford and Gurney, 2004; Donkin and Van Maanen, 2014). However, as shown in Harris et al. (2014), the RR_m decision rule automatically generates a range of optimum responses which follow a power law. Of our 4 tested rules the RR_m is the only decision rule that admits both a Piéron shape and a non-Piéron shape depending on α and q (see **Figure 2** and Appendix in Supplementary Materials). For the RR_m model, a Piéron's function is obtained only when $\alpha < q/(q+1)$. In a 2AFC, which is the classic paradigm used in studying with Piéron's law, $\alpha = 0.5$ which requires that $q > 1$; that is, the punishment must have a higher subjective magnitude than the reward for the function to follow Piéron's law. On the other hand, the non-Piéron's shape that we have observed for $\alpha = 0.75$ requires $\alpha > q/(q+1)$. This shape has never been observed in classic RT experiments. However, for a 2AFC experiment ($\alpha = 0.5$), it would require $q < 1$; that is, punishment would need have a *lower* subjective magnitude than reward. Thus, it is conceivable that the non-Piéron function may emerge in a 2AFC experiment, if the experimenter could augment the subjective reward over the punishment by manipulating the obtained gain or by verbal instruction. We are currently exploring this.

Time Series Analysis

The PACF showed that responses were dependent on the previous trials with a lag of one and weakly with a lag of two. We found that the change in response depended on whether the previous trial was rewarded or punished, and was particularly sensitive to punishment. A similar phenomenon is found in classic RT task, where trials following an error are usually substantially slower and more accurate than trial following a correct response (post-error slowing, e.g., Rabbitt, 1966; Laming, 1968; Brewer and Smith, 1984; Jentzsch and Dudschig, 2009; Strozyk and Jentzsch, 2012), and has been interpreted as a strategic adjustment of a response criterion (e.g., Brewer and Smith, 1984; Botvinick et al., 2001; Jentzsch and Leuthold, 2006). It is interesting to note that feedback does not give further accuracy information than the gauge, but may provide information on the rate of subjective reward (RR_m) which is not easily calculated from $ACC(t)$. It is possible that post error slowing arises the same mechanism, and may therefore reflect a reward rate maximizing algorithm.

Investigating the relation between the size of the adjustment and λ and α revealed the importance of examining $ACC(RT)$ as well as RT. RT clearly changed based on λ, but not on α. But $ACC(RT)$ depended on α but only weakly on λ. It is not possible to know exactly whether participants were adjusting their strategy based on RT or $ACC(RT)$, but the latter seems more likely. Is it conceivable that participants adjusted their accuracy by a certain proportion, and in so doing changed RT via its dependence on λ (Equation 1). The dependency on α may be explained by noticing that, when $\alpha = 0.25$, participants' range of adjustment is higher, because the gauge starts on a lower level, and therefore they have a larger range for adjustment than in the $\alpha = 0.75$ condition.

We also examined responses as a time series in order to explore possible strategies (or algorithms) used by the participants to maximize their subjective gain (e.g., exploration vs. exploitation). However, we found little evidence for a learning algorithm. The only significant trend between mean RT and condition duration was found for the condition $\alpha = 0.25$ and $\lambda = 0.16$. In this case, participants appeared to increase on average their response by 13.5 ms. We were not able to explain this trend, and why none of the other conditions showed a significant trend. Change in standard deviation could underpin a gradual switching from an exploratory to an exploitatory phase, as seen in reinforcement learning. However, we found no significant trends. This may suggest that participants are not using a learning algorithm with adjustable exploration factor, or that 3 minutes are not enough to detect any change in their strategy.

Normality in the Rate Domain

We have found that response distribution is not normal as seen in manual choice RT task (Harris et al., 2014) and saccades (Carpenter, 1981). This may be because participants are using a different decision mechanism than in a classic RT task. However, a different explanation can be offered. The reason for normality in the rate domain was explained as fluctuation in the relationship between response time and accuracy across trial, due to sensory noise in the stimulus (Harris et al., 2014). However, in our experiment this noise is drastically reduced, since $ACC(t)$ is always the same within a block. The response distribution could be a mixture of an optimization algorithm (see next section) and a fluctuation in $ACC(t)$. In classic RT experiment, the latter may usually mask the optimization algorithm. With the EXACT paradigm, it is possible to design an experiment in which $ACC(t)$ fluctuates across trials and test for normality in the rate domain, similarly to what predicted by Harris et al. (2014), but has yet to be tested.

Effect of Conditions on Distributions

We found that, in all conditions (except with $\alpha = 0.25$, $\lambda = 0.16$), the RT distributions were skewed on the right, similarly to what has been found in classic RT tasks. Also, increased mean resulted in increased the standard deviation of the distribution. These results are both observed in classic RT task (Luce, 1986). Within the framework of diffusion models, these results are explained by taking into account the geometry of the perceptual process. Most rise-to-threshold processes will in fact produce a right-skewed distribution which will generate less spread out distributions with more difficult conditions (higher mean RT; Ratcliff and Rouder, 1998). The explanation in terms of diffusion process does not seem to hold in this paradigm, since the participant is not accumulating information about a particular stimulus within a single trial, as usually assumed. We propose two alternative explanations for these results. Both explanations are for now speculative and clearly require more investigation. They both aim to describe what may be the underlying process that generates the distributions' shape observed without relying on a rise-to-threshold model.

Asymmetricity of the Decision Rule and Optimization Algorithm

The right-skewed distributions and the effect of different λ and α may be due to the way the participants search for the optimum t^*, and in particular may be due to the asymmetry of the decision rule functions (see **Figure 11**, red lines, for the RR_m decision rule, with a simulated dataset of responses). This might happen because responses that were too fast ($t < t^*$), would incur more cost than responses that were too slow ($t > t^*$) (for a given magnitude of error). Thus, any search strategy that was sensitive to the reduction in optimal reward rate would tend to err to the right side of the optimum t^*. Since the asymmetry of the decision rules depends on λ and α, this could also explain the relationship between distribution shape (e.g., standard deviation and skewness) and experimental condition found in our data (**Figure 8**). For example, when λ is small (slower gauge speed) or α is small the RR_m decision rule function becomes more symmetric and more flat, which implies that participants' gain in terms of RR_m would not change much, even with high response variability: the distribution would become less skewed and more spread out (compare the simulated distribution in **Figure 11** with our empirical results in **Figure 8**, upper left panel). On the other hand, as λ increases, the RR_m function becomes asymmetric, which would lead to a skewed and less disperse

distribution. This distribution would also be biased on the right side of t^* (red circles in **Figures 8**, **11**). In our data, most of the distributions follow for the RR_m decision rule, but not for the other decision rules. A peculiar exception is found with $\lambda = 0.16$ and $\alpha = 0.75$, but this could be partly due to the maximum point being close to zero, which would force the participant to produce a skewed distribution regardless of the smooth RR_m function.

This idea is an extension of an approach already presented by Bogacz et al. (2006) and Balci et al. (2011). In the latter work, empirical thresholds for the drift diffusion model were consistently found to be higher than the optimal threshold, which was deemed to be due to the asymmetry of the RR_mcurve. To our knowledge the same reasoning has not been applied to RT distributions or to the relationship between experimental conditions and distribution shapes.

Taking the asymmetry of the decision rule into account also means to slightly change the predicted t^* when fitting the data: all the prediction should be slightly slower than the one used now. We did not consider this, but suggest it as a possible future direction.

Accuracy is Normally Distributed

An alternative explanation could be that participants' responses are actually distributed such that their *accuracy* is normally distributed around $ACC(t^*)$, where t^* is the optimum point according to the decision rule used by the participants and the standard deviation can be interpreted as the individual precision parameter. $ACC\,(RT)$ cannot be analyzed in classic RT task because the accuracy of each single response is not known, whereas in the EXACT paradigm we can translate RT in $ACC\,(RT)$ by Equation (1), and ACC in $RT(ACC)$ by Equation (2).

Note that we should actually consider a truncated normal distribution in the accuracy domain (upper truncation is always one, and lower truncation is α in our experiment and 0.5 in 2AFC). The probability density is found to be equal to

$$\mathbf{RT}\left(t; t^*, \lambda, \alpha\right) = \frac{\phi\left(1 + (\alpha - 1)\exp\left(-\lambda t\right), ACC\left(t^*\right), \sigma\right)(1-\alpha)\,\lambda\exp\left(-\lambda t\right)}{\mathbf{\Phi}\left(1, ACC\left(t^*\right), \sigma\right) - \mathbf{\Phi}\left(\alpha, ACC\left(t^*\right), \sigma\right)} \quad (3)$$

Where ϕ is the normal probability function and $\mathbf{\Phi}$ is the cumulative normal probability function. **Figure 12** shows some examples of this distribution with different λ and α parameter. This generates an interesting relationship between the number of alternatives ($1/\alpha$), and the trial difficulty (λ). By increasing λ or α the distributions of RT become less spread out and the mean decreases, corresponding to faster responses. This is similar to what we empirically found with this paradigm and what is normally found in the classic RT task (Luce, 1986). This model does not take into account a sensory motor delay, but adding it would not drastically change the predictions.

In our dataset, most accuracy distributions were approximately Normal. However, this was not the case with small λ and high α (see **Figure 10**). Normality may be masked by severe truncation of the original datasets. In fact, as predicted by our model (**Figure 12**), the condition with small λ and high α contained the most severe truncated distributions. By excluding the most severely truncated distributions the remaining dataset appear less skewed but the deviation from Normality is still clear. This may be due to the small sample remaining after eliminating the seemly truncated distributions. It would be interesting to test whether accuracy of response distributions follow a near-Normal shape in classic RT tasks.

Limitations of the Exact Paradigm

The EXACT paradigm has some limitations and differences from the classic RT task. First, in the EXACT paradigm, the total time between the beginning of the trial and the participant's response

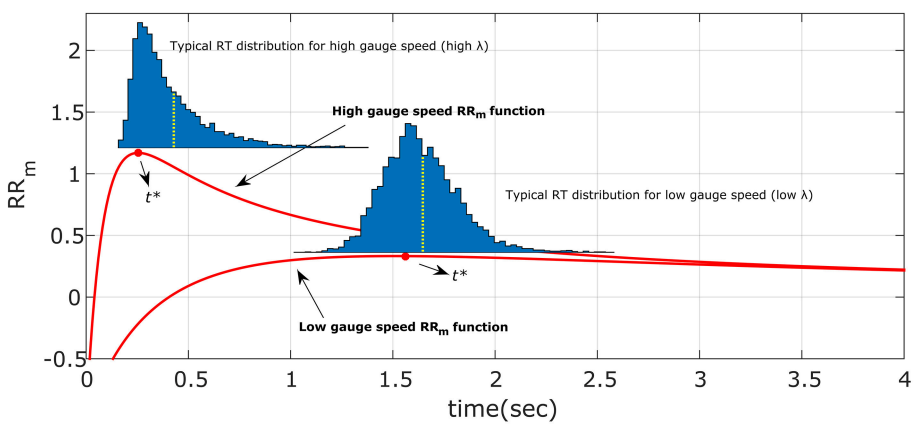


FIGURE 11 | Illustration of the hypothesized optimum distributions for the RR_m decision rule. The two RR_m functions (in red) for two different gauge speeds (different λ) are maximized by responding precisely at t^*. However, if the response is noisy, the asymmetry of RR_m make more profitable in terms of reward rate to err on the right side of t^* than on the left. The asymmetry effect is diminished with low gauge speed (small λ): the distributions are less skewed and more spread out. This also entails that the difference between distribution mean (yellow dotted lines) and t^* is greater with high gauge speed. Compare these simulated distributions with the empirical data in **Figure 8**.

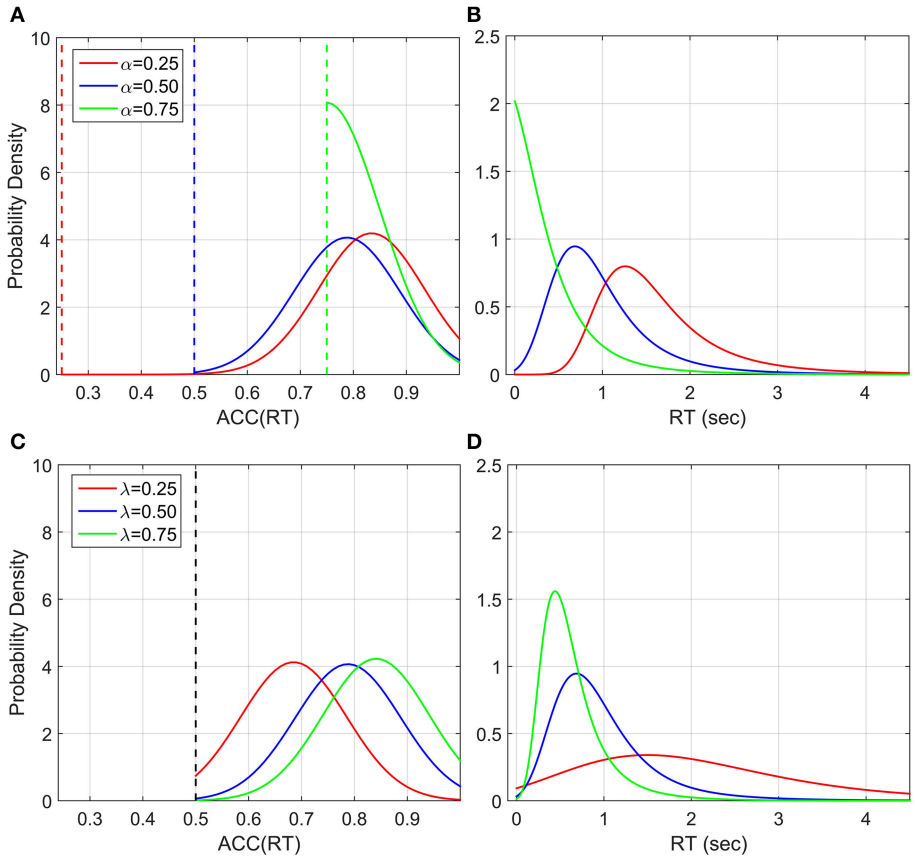


FIGURE 12 | Illustration of *ACC*(*RT*) distribution and corresponding RT distribution if ACC is normally distributed with truncation at 1 and at various α's. The mean of *ACC*(*RT*) is t^*_{RRm} and the standard deviation is set to 0.1 for illustration, which can be interpreted as a precision parameter. **(A)** *ACC*(*RT*) distributions for three α conditions with $\lambda = 1$. The dashed vertical lines correspond to the truncation point. **(B)** Corresponding RT distributions. By increasing α, which corresponds to decreasing the number of alternatives in a classic RT task, the distribution is less spread out and the mean decreases (faster responses). **(C)** *ACC*(*RT*) distributions for three λ conditions with $\alpha = 0.5$ and **(D)** corresponding RT distributions. By increasing λ, which corresponds to change the trial difficult in a classic RT task, the distribution become less spread and the mean decreases (faste responses).

corresponds to an increase in reward probability and includes the motor response time. It is not clear if participants take the motor response time into account and decide to respond accordingly. If not, desired accuracy would always be slightly lower than obtained accuracy. On other hand, in the classic RT experiment, most rise-to-threshold models assume motor response time to be an additional component to the accumulation time (Ratcliff and McKoon, 2008). We designed our experiment so that expected RT were of the order of seconds—much longer than typical motor response times so that any overestimation would be minimal. However, this remains to be further explored.

A second point is that the decision rule may differ between the EXACT and classic RT paradigms. Similarities in RT distribution, relationship between distribution shape and experimental condition, and the emerging of a Piéron's shape function lead us to deduce that participants are indeed using the same decision rule across different paradigms. However, the critical test would be to obtain the non-Piéron function in a low punishment classic RT experiment, which we are currently exploring. However, if this turned out not to be the case, it would still be interesting to understand why. Is the decision rule inherently entangled with the perceptual property of the task, or could it be due to other aspects of the task (different time scale, different instructions, etc.)?

The final limitation concerns the incentivisation scheme used. In our experiment we incentivized participants' with a prize for the overall best performance, but other schemes could be used. For example, earned points could be exchanged for money. We do not think this was important in the present experiment, since the amount of points earned/lost was the same for each condition, but it would be important in an experiment with reward/punishment manipulation. It is conceivable that different schemes may result in different decision rules, and future research could explore this possibility. Indeed, the EXACT paradigm may be a useful procedure for isolating the decision rule for different incentivisation schemes.

CONCLUSIONS

The analysis of the decision strategy used by humans has been constrained by the analysis of the perceptual process underlying

the decision itself. We designed an experiment, called the EXACT paradigm, which allows us to analyse participants' decision rules and responses based on the accuracy of their response. Instead of relying on a particular model based on a specific type of sampling process, the accumulator that substitutes the perceptual process is exogenously showed to the participant on a computer screen. This design allowed us to know in advance the relationship between response time and accuracy of response, $ACC(t)$. In this way it was possible to directly compare different decision rules, and found that RR_m provided the best fit. We suggest some innovative way to analyse the dataset that can be easily applied to classic RT tasks. Most importantly, we found relevant similarities between our distributions and distributions found in classic RT tasks: the distributions were generally skewed to the right and their shape depended on the trial difficulty. Two different models, one algorithmic and the other based on the accuracy of response, are proposed to explain the distributions shape and their dependency on experimental conditions. Both of these models establish a clear separation from the classic viewpoint of distributions as a result of a rise-to-threshold mechanism.

This new paradigm opens up new ways to explore the human decision-making process that are difficult or impossible using the classic RT paradigm including: exploring the behavior for unusual $ACC(t)$ functions (e.g., not-increasing, non-monotonic), easily manipulation of the number of alternatives (α), also by mimicking unusual setups ($\alpha > 0.5$), exploring subjective payoff (q), and exploring the "algorithm" used by participants use to find the optimum response time (t^*).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnbeh.2015.00288

REFERENCES

Balci, F., Simen, P., Niyogi, R., Saxe, A., Hughes, J., Holmes, P., et al. (2011). Acquisition of decision making criteria: reward rate ultimately beats accuracy. *Attent. Percept. Psychophys.* 73, 640–657. doi: 10.3758/s13414-010-0049-7

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765. doi: 10.1037/0033-295X.113.4.700

Bogacz, R., Hu, P. T., Holmes, P. J., and Cohen, J. D. (2010). Do humans produce the speed-accuracy trade-off that maximizes reward rate? *Q. J. Exp. Psychol. (Hove)* 63, 863–891. doi: 10.1080/17470210903091643

Bonnet, C., Zamora, M. C., Buratti, F., and Guirao, M. (1999). Group and individual gustatory reaction times and Pieron's Law. *Physiol. Behav.* 66, 549–558. doi: 10.1016/S0031-9384(98)00209-1

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi: 10.1037/0033-295X.108.3.624

Brewer, N., and Smith, G. A. (1984). How normal and retarded individuals monitor and regulate speed and accuracy of responding in serial choice tasks. *J. Exp. Psychol. Gen.* 113, 71–93. doi: 10.1037/0096-3445.113.1.71

Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cogn. Psychol.* 57, 153–178. doi: 10.1016/j.cogpsych.2007.12.002

Busemeyer, J. R., and Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol. Rev.* 100, 432–459. doi: 10.1037/0033-295X.100.3.432

Carpenter, R. H. S. (1981). "Oculomotor procrastination," in *Eye Movements: Cognition and Visual Perception*, eds D. F. Fisher, R. A. Monty, and J. W. Senders (Hillsdale: Lawrence Erlbaum), 237–246.

Donkin, C., and Van Maanen, L. (2014). Piéron's Law is not just an artifact of the response mechanism. *J. Math. Psychol.* 62–63, 22–32. doi: 10.1016/j.jmp.2014.09.006

Gold, J. I., and Shadlen, M. N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308. doi: 10.1016/S0896-6273(02)00971-6

Harris, C. M., Waddington, J., Biscione, V., and Manzi, S. (2014). Manual choice reaction times in the rate-domain. *Front. Hum. Neurosci.* 8:418. doi: 10.3389/fnhum.2014.00418

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Front. Neurosci.* 8:150. doi: 10.3389/fnins.2014.00150

Holmes, P., and Cohen, J. D. (2014). Optimality and some of its discontents: successes and shortcomings of existing models for binary decisions. *Top. Cogn. Sci.* 6, 258–278. doi: 10.1111/tops.12084

Jentzsch, I., and Dudschig, C. (2009). Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *Q. J. Exp. Psychol. (Hove)* 62, 209–218. doi: 10.1080/17470210802240655

Jentzsch, I., and Leuthold, H. (2006). Control over speeded actions: a common processing locus for micro- and macro-trade-offs? *Q. J. Exp. Psychol. (Hove)* 59, 1329–1337. doi: 10.1080/17470210600674394

Krajbich, I., and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13852–13857. doi: 10.1073/pnas.1101328108

LaBerge, D. (1962). A recruitment theory of simple behaviour. *Psychometrika* 27, 375–396. doi: 10.1007/BF02289645

Laming, D. R. J. (1968). *Information Theory of Choice-reaction Times*. New York, NY: Wiley.

Luce, R. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York, NY: Oxford University Press.

Maddox, W. T., and Bohil, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 1459–1482. doi: 10.1037/0278-7393.24.6.1459

Moran, O. (2014). Optimal decision making in heterogeneous and biased environments. *Psychon. Bull. Rev.* 22, 38–53. doi: 10.3758/s13423-014-0669-3

Piéron, H. (1914). Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices. *Année Psychol.* 22, 17–96.

Pins, D., and Bonnet, C. (1996). On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time. *Percept. Psychophys.* 58, 390–400. doi: 10.3758/BF03206815

Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *J. Exp. Psychol.* 71, 264–272. doi: 10.1037/h0022853

Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. doi: 10.1037/0033-295X.85.2.59

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychol. Bull.* 86, 446–461. doi: 10.1037/0033-2909.86.3.446

Ratcliff, R., Gomez, P., and McKoon, G. (2004a). A diffusion model account of the lexical decision task. *Psychol. Rev.* 111, 159–182. doi: 10.1037/0033-295X.111.1.159

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922. doi: 10.1162/neco.2008.12-06-420

Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356. doi: 10.1111/1467-9280.00067

Ratcliff, R., and Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *J. Exp. Psychol. Hum. Percep. Perform.* 26, 127–140. doi: 10.1037/0096-1523.26.1.127

Ratcliff, R., and Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* 111, 333–367. doi: 10.1037/0033-295X.111.2.333

Ratcliff, R., Thapar, A., and McKoon, G. (2004b). A diffusion model analysis of the effects of aging on recognition memory. *J. Mem. Lang.* 50, 408–424. doi: 10.1016/j.jml.2003.11.002

Ratcliff, R., Zandt, T. V., and McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychol. Rev.* 106, 261–300. doi: 10.1037/0033-295X.106.2.261

Simen, P., Balci, F., deSouza, L., Cohen, J. D., and Holmes, P. (2011). A model of interval timing by neural integration. *J. Neurosci.* 31, 9238–9253. doi: 10.1523/JNEUROSCI.3121-10.2011

Simen, P., and Cohen, J. D. (2009). Explicit melioration by a neural diffusion model. *Brain Res.* 1299, 95–117. doi: 10.1016/j.brainres.2009.07.017

Smith, P. L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168. doi: 10.1016/j.tins.2004.01.006

Stafford, T., and Gurney, K. N. (2004). The role of response mechanisms in determining reaction time performance: Piéron's law revisited. *Psychon. Bull. Rev.* 11, 975–987. doi: 10.3758/BF03196729

Stafford, T., Ingram, L., and Gurney, K. N. (2011). Piéron′s Law holds during stroop conflict: insights into the architecture of decision making. *Cogn. Sci.* 35, 1553–1566. doi: 10.1111/j.1551-6709.2011.01195.x

Stone, M. (1960). Models for choice-reaction time. *Psychometrika* 25, 251–260. doi: 10.1007/BF02289729

Strozyk, J. V., and Jentzsch, I. (2012). Weaker error signals do not reduce the effectiveness of post-error adjustments: comparing error processing in young and middle-aged adults. *Brain Res.* 1460, 41–49. doi: 10.1016/j.brainres.2012.04.028

Tversky, A., and Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertain.* 5, 297–323. doi: 10.1007/BF00122574

Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592. doi: 10.1037/0033-295X.108.3.550

Voss, A., Nagler, M., and Lerche, V. (2013). Diffusion models in experimental psychology: a practical introduction. *Exp. Psychol.* 60, 385–402. doi: 10.1027/1618-3169/a000218

Wald, A., and Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Stat.* 19, 326–339. doi: 10.1214/aoms/1177730197

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.